Invariance principles for homogeneous sums of free random variables

Aurélien Deya[1] and Ivan Nourdin[2]

Abstract: We extend, in the free probability framework, an invariance principle for multilinear homogeneous sums with low influences recently established by Mossel, O'Donnel and Oleszkiewicz in [6]. To do so, a hypercontractivity property for those homogeneous sums is necessary, and to prove it has turned out to be our main task. Finally, we deduce from our extension several universality phenomenons, in the spirit of the paper [10] by Nourdin, Peccati and Reinert.

Keywords: Central limit theorems; chaos; free Brownian motion; free probability; homogeneous sums; Lindeberg principle; universality, Wigner chaos.

AMS subject classifications: 46L54; 60H05; 60F05, 60F17

1. Introduction and background

Motivation and main goal. Our starting point is the following weak version (which is enough for our purpose) of an invariance principle for multilinear homogeneous sums with low influences, recently established in [6].

Theorem 1.1 (Mossel-O'Donnel-Oleszkiewicz). *Let $(\Omega, \mathcal{F}, P)$ be a probability space (in the classical sense). Let $X_1, X_2, \ldots$ (resp. $Y_1, Y_2, \ldots$) be a sequence of independent centered random variables with unit variance satisfying moreover*

$$\sup_{i\geq 1} E[|X_i|^r] < \infty \quad (\textit{resp. } \sup_{i\geq 1} E[|Y_i|^r] < \infty).$$

Fix $d \geq 1$, and consider a sequence of functions $f_N : \{1, \ldots, N\}^d \to \mathbb{R}$ satisfying the following two assumptions for each N and each $i_1, \ldots, i_d = 1, \ldots, N$:

(i) (full symmetry) $f_N(i_1, \ldots, i_d) = f_N(i_{\sigma(1)}, \ldots, i_{\sigma(d)})$ *for all* $\sigma \in \mathfrak{S}_d$*;*

(ii) (normalization) $d! \sum_{j_1,\ldots,j_d=1}^N f_N(j_1, \ldots, j_d)^2 = 1$.

Also, set

$$Q_N(x_1, \ldots, x_N) = \sum_{i_1,\ldots,i_d=1}^N f_N(i_1, \ldots, i_d) x_{i_1} \ldots x_{i_d} \tag{1}$$

and

$$\mathrm{Inf}_i(f_N) = \sum_{j_2,\ldots,j_d=1}^N f_N(i, j_2, \ldots, j_d)^2, \quad i = 1, \ldots, N.$$

Then, for any integer $m \geq 1$,

$$E[Q_N(X_1, \ldots, X_N)^m] - E[Q_N(Y_1, \ldots, Y_N)^m] = O(\tau_N^{1/2}), \tag{2}$$

[1]Institut Élie Cartan, Université Henri Poincaré, BP 70239, 54506 Vandoeuvre-lès-Nancy, France. Email: `Aurelien.Deya@iecn.u-nancy.fr`

[2]Institut Élie Cartan, Université Henri Poincaré, BP 70239, 54506 Vandoeuvre-lès-Nancy, France. Email: `inourdin@gmail.com`. Supported in part by the two following (french) ANR grants: 'Exploration des Chemins Rugueux' [ANR-09-BLAN-0114] and 'Malliavin, Stein and Stochastic Equations with Irregular Coefficients' [ANR-10-BLAN-0121].

where $\tau_N = \max_{1\le i\le N} \mathrm{Inf}_i(f_N)$.

In [6], the authors were motivated by solving two conjectures, namely the *Majority Is Stablest* conjecture from theoretical computer science and the *It Ain't Over Till It's Over* conjecture from social choice theory. It is worthwhile noting that there is another striking consequence of Theorem 1.1, more in the spirit of the classical central limit theorem. Indeed, in article [10] Nourdin, Peccati and Reinert combined Theorem 1.1 with the celebrated *Fourth Moment Theorem* of Nualart and Peccati [11], and deduced that multilinear homogenous sums of general centered independent random variables with unit variance enjoy the following universality phenomenon.

Theorem 1.2 (Nourdin-Peccati-Reinert). *Let $(\Omega, \mathcal{F}, P)$ be a probability space (in the classical sense). Let $G_1, G_2, \ldots$ be a sequence of i.i.d. $\mathcal{N}(0,1)$ random variables. Fix $d \ge 2$ and consider a sequence of functions $f_N : \{1,\ldots,N\}^d \to \mathbb{R}$ satisfying the following three assumptions for each N and each $i_1,\ldots,i_d = 1,\ldots,N$:*

(i) (full symmetry) $f_N(i_1,\ldots,i_d) = f_N(i_{\sigma(1)},\ldots,i_{\sigma(d)})$ *for all* $\sigma \in \mathfrak{S}_d$;
(ii) (vanishing on diagonals) $f_N(i_1,\ldots,i_d) = 0$ *if* $i_k = i_l$ *for some* $k \ne l$;
(iii) (normalization) $d! \sum_{j_1,\ldots,j_d=1}^{N} f_N(j_1,\ldots,j_d)^2 = 1$.

Also, let $Q_N(x_1,\ldots,x_N)$ be given by (1). Then, the following two conclusions are equivalent as $N \to \infty$:

(A) $Q_N(G_1,\ldots,G_N) \overset{\text{law}}{\to} \mathcal{N}(0,1)$;
(B) $Q_N(X_1,\ldots,X_N) \overset{\text{law}}{\to} \mathcal{N}(0,1)$ *for* any *sequence $X_1, X_2, \ldots$ of i.i.d. centered random variables with unit variance and all moments.*

In the present paper, our goal is twofold. We shall first extend Theorem 1.1 in the context of free probability and we shall then investigate whether a result such as Theorem 1.2 continues to hold true in this framework. We are motivated by the fact that there is often a close correspondence between classical probability and free probability, in which the Gaussian law (resp. the classical notion of independence) has the semicircular law (resp. the notion of free independence) as an analogue.

Free probability in a nutshell. Before going into details and for the sake of clarity, let us first introduce some of the central concepts in the theory of free probability. (See [8] for a systematic presentation.)

A *free tracial probability space* is a von Neumann algebra $\mathcal{A}$ (that is, an algebra of operators on a real separable Hilbert space, closed under adjoint and convergence in the weak operator topology) equipped with a *trace* φ, that is, a unital linear functional (meaning preserving the identity) which is weakly continuous, positive (meaning $\varphi(X) \ge 0$ whenever X is a non-negative element of $\mathcal{A}$; i.e. whenever $X = YY^*$ for some $Y \in \mathcal{A}$), faithful (meaning that if $\varphi(YY^*) = 0$ then $Y = 0$), and tracial (meaning that $\varphi(XY) = \varphi(YX)$ for all $X, Y \in \mathcal{A}$, even though in general $XY \ne YX$).

In a free tracial probability space, we refer to the self-adjoint elements of the algebra as *random variables*. Any random variable X has a *law*: this is the unique probability measure μ on $\mathbb{R}$ with the same moments as X; in other words, μ is such that

$$\int_{\mathbb{R}} Q(x)d\mu(x) = \varphi(Q(X)), \tag{3}$$

for any real polynomial Q.

In the free probability setting, the notion of *independence* (introduced by Voiculescu in [13]) goes as follows. Let $\mathcal{A}_1, \ldots, \mathcal{A}_p$ be unital subalgebras of $\mathcal{A}$. Let $X_1, \ldots, X_m$ be elements chosen among the $\mathcal{A}_i$'s such that, for $1 \le j < m$, two consecutive elements X_j and X_{j+1} do not come

from the same $\mathcal{A}_i$, and such that $\varphi(X_j) = 0$ for each j. The subalgebras $\mathcal{A}_1, \ldots, \mathcal{A}_p$ are said to be *free* or *freely independent* if, in this circumstance,

$$\varphi(X_1 X_2 \cdots X_m) = 0. \tag{4}$$

Random variables are called freely independent if the unital algebras they generate are freely independent. If X, Y are freely independent, then their joint moments are determined by the moments of X and Y separately as in the classical case.

The *semicircular distribution* $\mathcal{S}(m, \sigma^2)$ with mean $m \in \mathbb{R}$ and variance $\sigma^2 > 0$ is the probability distribution

$$\mathcal{S}(m, \sigma^2)(dx) = \frac{1}{2\pi\sigma^2}\sqrt{4\sigma^2 - (x-m)^2}\, \mathbf{1}_{\{|x-m| \le 2\sigma\}}\, dx.$$

If $m = 0$, this distribution is symmetric around 0, and therefore its odd moments are all 0. A simple calculation shows that the even centered moments are given by (scaled) Catalan numbers: for non-negative integers k,

$$\int_{m-2\sigma}^{m+2\sigma} (x-m)^{2k} \mathcal{S}(m, \sigma^2)(dx) = C_k \sigma^{2k},$$

where $C_k = \frac{1}{k+1}\binom{2k}{k}$ (see, e.g., [8, Lecture 2]).

Our main results. We are now in a position to state our first main result, which is nothing but a suitable generalization of Theorem 1.1 in the free probability setting.

Theorem 1.3. *Let $(\mathcal{A}, \varphi)$ be a free tracial probability space. Let $X_1, X_2, \ldots$ (resp. $Y_1, Y_2, \ldots$) be a sequence of centered free random variables with unit variance (that is, such that $\varphi(X_i^2) = \varphi(Y_i^2) = 1$ for all i), satisfying moreover*

$$\sup_{i \ge 1} \varphi(|X_i|^r) < \infty \quad \textit{(resp. } \sup_{i \ge 1} \varphi(|Y_i|^r) < \infty\textit{)} \quad \textit{for all } r \ge 1,$$

where $|X| = \sqrt{X^*X}$. Fix $d \ge 1$, and consider a sequence of functions $f_N : \{1, \ldots, N\}^d \to \mathbb{R}$ satisfying the following three assumptions for each N and each $i_1, \ldots, i_d = 1, \ldots, N$:

(i) (mirror-symmetry) $f_N(i_1, \ldots, i_d) = f_N(i_d, \ldots, i_1)$;

(ii) (vanishing on diagonals) $f_N(i_1, \ldots, i_d) = 0$ *if* $i_k = i_l$ *for some* $k \ne l$;

(iii) (normalization) $\sum_{j_1, \ldots, j_d = 1}^N f_N(j_1, \ldots, j_d)^2 = 1$.

Also, set

$$Q_N(x_1, \ldots, x_N) = \sum_{i_1, \ldots, i_d = 1}^N f_N(i_1, \ldots, i_d) x_{i_1} \ldots x_{i_d} \tag{5}$$

and

$$\mathrm{Inf}_i(f_N) = \sum_{l=1}^d \sum_{j_1, \ldots, j_{d-1} = 1}^N f_N(j_1, \ldots, j_{l-1}, i, j_l, \ldots, j_{d-1})^2, \quad i = 1, \ldots, N.$$

Then, for any integer $m \ge 1$,

$$\varphi(Q_N(X_1, \ldots, X_N)^m) - \varphi(Q_N(Y_1, \ldots, Y_N)^m) = O(\tau_N^{1/2}), \tag{6}$$

where $\tau_N = \max_{1 \le i \le N} \mathrm{Inf}_i(f_N)$.

Due to the lack of commutativity in the free context, the proof of Theorem 1.3 is far more complicated than its commutative counterpart. Moreover, it is worthwhile noting that it contains the free central limit theorem as an immediate corollary. Indeed, let us choose $d = 1$ (in this case, assumptions (i) and (ii) are of course immaterial), $Y_1, Y_2, \ldots \sim \mathcal{S}(0,1)$ and $f_N(i) = \frac{1}{\sqrt{N}}$, $i = 1, \ldots, N$. We then have $Q_N(Y_1, \ldots, Y_N) \sim \mathcal{S}(0,1) \overset{\text{law}}{=} Y_1$ (thanks to (iii) as well as the

fact that a sum of freely independent semicircular random variables remains semicircular) and $\tau_N \to 0$ as $N \to \infty$, so that, thanks to (6),

$$\varphi\left[\left(\frac{X_1 + \ldots + X_N}{\sqrt{N}}\right)^m\right] \to \varphi(Y_1^m)$$

for each $m \geq 1$ as $N \to \infty$, which is exactly what the free central limit theorem asserts.

When $d \geq 2$, by combining Theorem 1.3 with the main finding of [4], we will prove the following free counterpart of Theorem 1.2.

Theorem 1.4. *Let $(\mathcal{A}, \varphi)$ be a free tracial probability space. Let $S_1, S_2, \ldots$ be a sequence of free $\mathcal{S}(0,1)$ random variables. Fix $d \geq 2$ and consider a sequence of functions $f_N : \{1, \ldots, N\}^d \to \mathbb{R}$ satisfying the following three assumptions for each N and each $i_1, \ldots, i_d = 1, \ldots, N$:*

- *(i)* (full symmetry) $f_N(i_1, \ldots, i_d) = f_N(i_{\sigma(1)}, \ldots, i_{\sigma(d)})$ *for all* $\sigma \in \mathfrak{S}_d$*;*
- *(ii)* (vanishing on diagonals) $f_N(i_1, \ldots, i_d) = 0$ *if* $i_k = i_l$ *for some* $k \neq l$*;*
- *(iii)* (normalization) $\sum_{j_1,\ldots,j_d=1}^{N} f_N(j_1, \ldots, j_d)^2 = 1$.

Also, let $Q_N(x_1, \ldots, x_N)$ be the polynomial in non-commuting variables given by (5). Then, the following two conclusions are equivalent as $N \to \infty$:

- *(A)* $Q_N(S_1, \ldots, S_N) \overset{\text{law}}{\to} \mathcal{S}(0,1)$*;*
- *(B)* $Q_N(X_1, \ldots, X_N) \overset{\text{law}}{\to} \mathcal{S}(0,1)$ *for any sequence $X_1, X_2, \ldots$ of free identically distributed and centered random variables with unit variance.*

Although a weak 'mirror-symmetry' assumption would have been undoubtedly more natural, we impose in Theorem 1.4 the same 'full symmetry' assumption (i) than in Theorem 1.2. This is unfortunately not insignificant in our non-commutative framework. But we cannot expect better by using our strategy of proof, as is illustrated by a concrete counterexample in Section 2.

Theorem 1.4 may be seen as a free universality phenomenon, in the sense that the semicircular behavior of $Q_N(X_1, \ldots, X_N)$ is asymptotically insensitive to the distribution of its summands. In reality, this is more subtle, as the following explicit situation well illustrates in the case $d = 2$ (quadratic case). Indeed, let us consider

$$Q_N(x_1, \ldots, x_N) = \frac{1}{\sqrt{2N-2}} \sum_{i=2}^{N} (x_1 x_i + x_i x_1), \quad N \geq 2,$$

let $S_1, S_2, \ldots$ be a sequence of free $\mathcal{S}(0,1)$ random variables and let $X_1, X_2, \ldots$ be a sequence of free Rademacher random variables (that is, the law of X_1 is given by $\frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}$). Then $Q_N(X_1, \ldots, X_N) \overset{\text{law}}{\to} \mathcal{S}(0,1)$ as $N \to \infty$, but

$$Q_N(S_1, \ldots, S_N) \overset{\text{law}}{\to} \frac{1}{\sqrt{2}}(S_1 S_2 + S_2 S_1) \not\sim \mathcal{S}(0,1).$$

(See Section 2 for the details.) This means that it is possible to have $Q_N(X_1, \ldots, X_N)$ converging in law to $\mathcal{S}(0,1)$ for a *particular* centered distribution of X_1, without having the same phenomenon for *every* centered distribution with variance one. The question of which are the distributions that enjoy such a universality phenomenon is still an open problem. (In the commutative case, it is known that the Gaussian and the Poisson distributions both lead to universality, see [10, 12]. Yet there are no other examples.)

Organization of the paper. The rest of our paper is organized as follows. In Section 2, we deduce from Theorem 1.3 several results connected with the universality phenomenon and we study the limitations of Theorem 1.4. Section 3 is devoted to the proof of Theorem 1.3, which turns out to be the main task of this paper. The proof of a technical estimate, needed in the proof of Theorem 1.3, is postponed in Section 4 for the sake of clarity.

2. Free universality

In this section, we show how Theorem 1.3 leads to several results connected with the universality phenomenon. We also study the limitations of Theorem 1.4: Can we replace the role played by the semicircular distribution by any other law? Can we replace the full symmetry assumption (i) by a more natural one?

To do so, we first need to recall some facts proven in references [1, 4].

Convergence of Wigner integrals. For $1 \leq p \leq \infty$, we write $L^p(\mathcal{A}, \varphi)$ to indicate the L^p space obtained as the completion of $\mathcal{A}$ with respect to the norm $\|A\|_p = \varphi(|A|^p)^{1/p}$, where $|A| = \sqrt{A^*A}$, and $\|\cdot\|_\infty$ stands for the operator norm. For every integer $q \geq 2$, the space $L^2(\mathbb{R}_+^q)$ is the collection of all real-valued functions on $\mathbb{R}_+^q$ that are square-integrable with respect to the Lebesgue measure. Given $f \in L^2(\mathbb{R}_+^q)$, we write $f^*(t_1, t_2, ..., t_q) = f(t_q, ..., t_2, t_1)$, and we call f^* the *adjoint* of f. We say that an element of $L^2(\mathbb{R}_+^q)$ is *mirror symmetric* whenever $f = f^*$ as a function. Given $f \in L^2(\mathbb{R}_+^q)$ and $g \in L^2(\mathbb{R}_+^p)$, for every $r = 1, ..., p \wedge q$ we define the rth *contraction* of f and g as the element of $L^2(\mathbb{R}_+^{p+q-2r})$ given by

$$\begin{aligned} & f \overset{r}{\frown} g(t_1, \ldots, t_{p+q-2r}) \\ = & \int_{\mathbb{R}_+^{p+q-2r}} f(t_1, \ldots, t_{p-r}, x_1, \ldots, x_r) g(x_r, \ldots, x_1, t_{p-r+1}, \ldots, t_{p+q-2r}) dx_1 \ldots dx_r. \end{aligned} \tag{7}$$

One also writes $f \overset{0}{\frown} g(t_1, ..., t_{p+q}) = f \otimes g(t_1, ..., t_{p+q}) = f(t_1, ..., t_q) g(t_{q+1}, ..., t_{p+q})$. In the following, we shall use the notations $f \overset{0}{\frown} g$ and $f \otimes g$ interchangeably. Observe that, if $p = q$, then $f \overset{p}{\frown} g = \langle f, g^* \rangle_{L^2(\mathbb{R}_+^q)}$.

A *free Brownian motion* S on $(\mathcal{A}, \varphi)$ consists of: (i) a filtration $\{\mathcal{A}_t : t \geq 0\}$ of von Neumann sub-algebras of $\mathcal{A}$ (in particular, $\mathcal{A}_u \subset \mathcal{A}_t$ for $0 \leq u < t$), (ii) a collection $S = (S_t)_{t \geq 0}$ of self-adjoint operators such that:

- $S_t \in \mathcal{A}_t$ for every t;
- for every t, S_t has a semicircular distribution $\mathcal{S}(0, t)$;
- for every $0 \leq u < t$, the increment $S_t - S_u$ is freely independent of $\mathcal{A}_u$, and has a semicircular distribution $\mathcal{S}(0, t - u)$.

For every integer $q \geq 1$, the collection of all random variables of the type $I_q(f)$, $f \in L^2(\mathbb{R}_+^q)$, is called the qth *Wigner chaos* associated with S, and is defined according to [1, Section 5.3], namely:

- first define $I_q(f) = (S_{b_1} - S_{a_1}) \ldots (S_{b_q} - S_{a_q})$ for every function f having the form
$$f(t_1, ..., t_q) = \mathbf{1}_{(a_1, b_1)}(t_1) \times \ldots \times \mathbf{1}_{(a_q, b_q)}(t_q), \tag{8}$$
where the intervals (a_i, b_i), $i = 1, ..., q$, are pairwise disjoint;
- extend linearly the definition of $I_q(f)$ to simple functions vanishing on diagonals, that is, to functions f that are finite linear combinations of indicators of the type (8);
- exploit the isometric relation
$$\langle I_q(f_1), I_q(f_2) \rangle_{L^2(\mathcal{A}, \varphi)} = \varphi\left(I_q(f_1)^* I_q(f_2)\right) = \varphi\left(I_q(f_1^*) I_q(f_2)\right) = \langle f_1, f_2 \rangle_{L^2(\mathbb{R}_+^q)}, \tag{9}$$
where f_1, f_2 are simple functions vanishing on diagonals, and use a density argument to define $I_q(f)$ for a general $f \in L^2(\mathbb{R}_+^q)$.

Observe that relation (9) continues to hold for every pair $f_1, f_2 \in L^2(\mathbb{R}_+^q)$. Moreover, the above sketched construction implies that $I_q(f)$ is self-adjoint if and only if f is mirror symmetric. We recall the following fundamental multiplication formula, proven in [1]. For every $f \in L^2(\mathbb{R}_+^p)$

and $g \in L^2(\mathbb{R}_+^q)$, where $p, q \geq 1$, we have

$$I_p(f)I_q(g) = \sum_{r=0}^{p\wedge q} I_{p+q-2r}(f \overset{r}{\frown} g). \tag{10}$$

Let $S_1, S_2, \ldots \sim \mathcal{S}(0,1)$ be freely independent, fix $d \geq 2$, and consider a sequence of functions $f_N : \{1,\ldots,N\}^d \to \mathbb{R}$ satisfying assumption (ii) and (iii) of Theorem 1.4 as well as

$$f_N(i_1,\ldots,i_d) = f_N(i_d,\ldots,i_1) \quad \text{for all } N \geq 1 \text{ and } i_1,\ldots,i_d \in \{1,\ldots,N\}. \tag{11}$$

Let also $Q_N(x_1,\ldots,x_N)$ be the polynomial in non-commuting variables given by (5). Set $e_i = \mathbf{1}_{[i-1,i]} \in L^2(\mathbb{R}_+)$, $i \geq 1$. For each N, one has

$$Q_N(S_1,\ldots,S_N) \overset{\text{law}}{=} Q_N(I_1(e_1),\ldots,I_1(e_N)). \tag{12}$$

By applying the multiplication formula (10) and by taking into account assumption (ii), it is straightforward to check that

$$Q_N(I_1(e_1),\ldots,I_1(e_N)) = I_d(g_N), \tag{13}$$

where

$$g_N = \sum_{i_1,\ldots,i_d=1}^{N} f_N(i_1,\ldots,i_d) e_{i_1} \otimes \ldots \otimes e_{i_d}. \tag{14}$$

The function g_N is mirror-symmetric (due to (11)) and has an $L^2(\mathbb{R}_+^d)$-norm equal to 1 (due to (iii)). Using both Theorems 1.3 and 1.6 of [4] (see also [9]), we deduce that the following equivalence holds true as $N \to \infty$:

$$Q_N(S_1,\ldots,S_N) \overset{\text{law}}{\to} \mathcal{S}(0,1) \iff \|g_N \overset{r}{\frown} g_N\|_{L^2(\mathbb{R}_+^{2d-2r})} \to 0 \;\text{ for all } r \in \{1,\ldots,d-1\}. \tag{15}$$

For $r = d-1$, observe that

$$\begin{aligned}
\|g_N \overset{d-1}{\frown} g_N\|_{L^2(\mathbb{R}_+^2)} &= \left\| \sum_{i,j=1}^{N} \left(\sum_{k_2,\ldots,k_d=1}^{N} f_N(i,k_2,\ldots,k_d) f_N(k_d,\ldots,k_2,j) \right) e_i \otimes e_j. \right\|_{L^2(\mathbb{R}_+^2)} \\
&= \sqrt{\sum_{i,j=1}^{N} \left(\sum_{k_2,\ldots,k_d=1}^{N} f_N(i,k_2,\ldots,k_d) f_N(k_d,\ldots,k_2,j) \right)^2} \\
&\geq \sqrt{\sum_{i=1}^{N} \left(\sum_{k_2,\ldots,k_d=1}^{N} f_N(i,k_2,\ldots,k_d)^2 \right)^2} \quad \text{(by setting } j = i \text{ and using (11))} \\
&\geq \max_{i=1,\ldots,N} \sum_{k_2,\ldots,k_d=1}^{N} f_N(i,k_2,\ldots,k_d)^2.
\end{aligned} \tag{16}$$

Proof of Theorem 1.4. Of course, only the implication $(A) \to (B)$ must be shown. Assume that (A) holds. Then, using (15) (condition (i) implies in particular (11)), we get that $\|g_N \overset{d-1}{\frown} g_N\|_{L^2(\mathbb{R}_+^2)} \to 0$ as $N \to \infty$. Using (16) and since f_N is fully-symmetric, we deduce that the quantity τ_N of Theorem 1.3 tends to zero as N goes to infinity. This, combined with assumption (A) and (6), leads to (B). $\square$

A counterexample. In Theorem 1.4, can we replace the role played by the semicircular distribution by any other law? The answer is no in general. Indeed, let us take a look at the following situation. Fix $d = 2$ and consider

$$Q_N(x_1,\dots,x_N) = \frac{1}{\sqrt{2N-2}}\sum_{i=2}^{N}(x_1x_i + x_ix_1), \quad N \geq 2.$$

Let $S_1, S_2, \dots$ be a sequence of free $\mathcal{S}(0,1)$ random variables and let $X_1, X_2, \dots$ be a sequence of free Rademacher random variables (that is, the law of X_1 is given by $\frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}$). Then, using the free central limit theorem, it is clear on one hand that

$$\begin{aligned} Q_N(X_1,\dots,X_N) &= \frac{1}{\sqrt{2}}X_1\left(\frac{1}{\sqrt{N-1}}\sum_{i=2}^{N}X_i\right) + \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{N-1}}\sum_{i=2}^{N}X_i\right)X_1 \\ &\overset{\text{law}}{\to} \frac{1}{\sqrt{2}}\big(X_1S_1 + S_1X_1\big) \quad \text{as } N\to\infty, \end{aligned}$$

with X_1 and S_1 freely independent. By Proposition 1.10 and identity (1.10) of Nica and Speicher [7], it turns out that $\frac{1}{\sqrt{2}}\big(X_1S_1 + S_1X_1\big) \sim \mathcal{S}(0,1)$. But, on the other hand,

$$\begin{aligned} Q_N(S_1,\dots,S_N) &= \frac{1}{\sqrt{2}}S_1\left(\frac{1}{\sqrt{N-1}}\sum_{i=2}^{N}S_i\right) + \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{N-1}}\sum_{i=2}^{N}S_i\right)S_1 \\ &\overset{\text{law}}{=} \frac{1}{\sqrt{2}}\big(S_1S_2 + S_2S_1\big). \end{aligned}$$

The random variable $\frac{1}{\sqrt{2}}\big(S_1S_2 + S_2S_1\big)$ being *not* $\mathcal{S}(0,1)$ distributed (its law is indeed the so-called *tetilla law*, see [2]), we deduce that one cannot replace the role played by the semicircular distribution in Theorem 1.4 by the Rademacher distribution.

Another counterexample. In Theorem 1.4, can we replace the full symmetry assumption (i) by the mirror-symmetry assumption? Unfortunately, we have not been able to answer this question. But if the answer is yes, what is sure is that we cannot use the same arguments as in the fully-symmetric case to show such a result. Indeed, when f_N is fully-symmetric we have

$$\tau_N = d \times \max_{i=1,\dots,N}\sum_{k_2,\dots,k_d=1}^{N} f_N(i,k_2,\dots,k_d)^2,$$

allowing us to prove Theorem 1.4 by using the following set of implications: as $N\to\infty$,

$$\begin{aligned} Q_N(S_1,\dots,S_N) \overset{\text{law}}{\to} \mathcal{S}(0,1) \quad &\overset{(15)}{\Longrightarrow} \quad \|g_N \overset{d-1}{\frown} g_N\|_{L^2(\mathbb{R}_+^2)} \overset{(16)}{\Longrightarrow} \tau_N \to 0 \\ &\overset{\text{Thm 1.3}}{\Longrightarrow} \quad Q_N(X_1,\dots,X_N) \overset{\text{law}}{\to} \mathcal{S}(0,1). \end{aligned} \tag{17}$$

Unfortunately, when f_N is only mirror-symmetric the implication

$$\|g_N \overset{d-1}{\frown} g_N\|_{L^2(\mathbb{R}_+^2)} \Longrightarrow \tau_N \to 0, \tag{18}$$

that plays a crucial role in (17), is no longer true in general. To see why, let us consider the following counterexample (for which we fix $d = 3$). Define first a sequence of functions $f'_N : \{1,\dots,N\}^2 \to \mathbb{R}$ according to the formula

$$f'_N(i,i+1) = f'_N(i+1,i) = \frac{1}{\sqrt{2N-2}},$$

and $f'_N(i,j)=0$ whenever $i=j$ or $|j-i|\geq 2$. Next, for $i,j,k\in\{1,\ldots,N\}$, set

$$f_N(i,j,k)=\begin{cases} 0 & \text{if } j\geq 2 \text{ or } (j=1 \text{ and } i=1) \text{ or } (j=1 \text{ and } k=1) \\ f'_{N-1}(i-1,k-1) & \text{otherwise.}\end{cases} \tag{19}$$

Easy-to-check properties of f_N include mirror-symmetry, vanishing on diagonals property,

$$\sum_{i,j,k=1}^{N} f_N(i,j,k)^2=\sum_{i,k=1}^{N-1} f'_{N-1}(i,k)^2=1$$

and

$$\sum_{i,j=1}^{N}\left(\sum_{k,l=1}^{N} f_N(i,k,l)f_N(l,k,j)\right)^2=\sum_{i,j=1}^{N}\left(\sum_{l=1}^{N-1} f'_{N-1}(i,l)f'_{N-1}(l,j)\right)^2\to 0. \tag{20}$$

Let g_N be given by (14), that is,

$$g_N=\frac{1}{\sqrt{2N-4}}\sum_{i=1}^{N-2}\big(e_{i+1}\otimes e_1\otimes e_{i+2}+e_{i+2}\otimes e_1\otimes e_{i+1}\big).$$

The limit (20) can be readily translated into $\|g_N\overset{2}{\frown}g_N\|^2_{L^2(\mathbb{R}_+^2)}\to 0$ as $N\to\infty$. On the other hand, we have

$$\begin{aligned}\tau_N=\max_{1\leq j\leq N}\mathrm{Inf}_j(f_N) &= \max_{1\leq j\leq N}\sum_{i,k=1}^{N}\{f_N(i,j,k)^2+f_N(j,i,k)^2+f_N(i,k,j)^2\}\\ &\geq \max_{1\leq j\leq N}\sum_{i,k=1}^{N} f_N(i,j,k)^2 = \sum_{i,k=1}^{N} f_N(i,1,k)^2 = 1,\end{aligned}$$

which contradicts (18), as announced.

It is also worth noting that the sequence of functions f_N defined by (19) provides an explicit counterexample to the so-called *Wiener-Wigner transfer principle* (see [4, Theorem 1.8]) in a non fully-symmetric situation. Indeed, on one hand, we have

$$\|g_N\overset{1}{\frown}g_N\|^2_{L^2(\mathbb{R}_+^2)}=\|g_N\overset{2}{\frown}g_N\|^2_{L^2(\mathbb{R}_+^2)}\to 0\quad\text{as } N\to\infty,$$

which, due to (15), entails that $Q_N(S_1,\ldots,S_N)\overset{\text{law}}{\to}\mathcal{S}(0,1)$. On the other hand, let $G_1,\ldots,G_N\sim\mathcal{N}(0,1)$ be independent random variables defined on a (classical) probability space $(\Omega,\mathcal{F},P)$. One has

$$Q_N(G_1,\ldots,G_N)=G_1\times\left(\frac{2}{\sqrt{2N-4}}\sum_{i=2}^{N-1}G_iG_{i+1}\right),$$

and it is easily checked that $\frac{2}{\sqrt{2N-4}}\sum_{i=2}^{N-1}G_iG_{i+1}\overset{\text{law}}{\to}\mathcal{N}(0,2)$ (apply, e.g., the Fourth Moment Theorem of [11]). As a result, the sequence $Q_N(G_1,\ldots,G_N)$ converges in law to $\sqrt{2}\,G_1G_2$, which is not Gaussian. This leads to our desired contradiction.

Free CLT for homogeneous sums. As an application of Theorem 1.3, let us also highlight the following practical convergence criterion for multilinear polynomials, which can be readily derived from (15).

Theorem 2.1. *Let $(\mathcal{A},\varphi)$ be a free tracial probability space. Let $X_1,X_2,\ldots$ be a sequence of centered free random variables with unit variance satisfying $\sup_{i\geq 1}\varphi(|X_i|^r)<\infty$ for all $r\geq 1$. Fix $d\geq 1$, and consider a sequence of functions $f_N:\{1,\ldots,N\}^d\to\mathbb{R}$ satisfying the three basic assumptions (i)-(ii)-(iii) of Theorem 1.3. Assume moreover that, as N tends to infinity,*

$\max_{1\le j\le N} \mathrm{Inf}_j(f_N) \to 0$ *and* $\|g_N \overset{r}{\frown} g_N\|_{L^2(\mathbb{R}_+^{2d-2r})} \to 0$ *for all* $r \in \{1,\dots,d-1\}$*, where* g_N *is defined through (14). Then one has*

$$\sum_{i_1,\dots,i_d=1}^{N} f_N(i_1,\dots,i_d) X_{i_1}\dots X_{i_d} \overset{\text{law}}{\to} \mathcal{S}(0,1). \tag{21}$$

For instance, thanks to this result one can easily check that, given a positive integer k, one has

$$\frac{1}{\sqrt{N}} \sum_{i=1}^{N-k} \{X_i X_{i+1}\dots X_{i+k} + X_{i+k}X_{i+k-1}\dots X_i\} \overset{\text{law}}{\to} \mathcal{S}(0,1) \quad \text{as } N\to\infty$$

for any sequence (X_i) of centered free random variables with unit variance satisfying $\sup_{i\ge1}\varphi(|X_i|^r) < \infty$ for all $r \ge 1$.

3. Proof of Theorem 1.3

As in [6], our strategy is essentially based on a generalization of the classical Lindeberg method, which was originally designed for linear sums of (classical) random variables (see [5]). Before we turn to the details of the proof, let us briefly report the two main differences with the arguments displayed in [6] for commuting random variables.

First, in this non-commutative context, we can no longer rely on some classical Taylor expansion as a starting point of our study. This issue can be easily overcome though, by resorting to abstract expansion formulae (see (24)) together with appropriate Hölder-type estimates (see (28)). As far as this particular point is concerned, the situation is quite similar to what can be found in [3], even if the latter reference is only concerned with the linear case, i.e., $d = 1$.

In fact, the main additional difficulty raised by this *free* background lies in the transposition of the hypercontractivity property, which is at the core of the procedure. Indeed, it appears quickly in [6] that the proof of hypercontractivity for multilinear polynomials heavily depends on the fact that the variables do commute (see, e.g., the proof of [6, Proposition 3.11]), so that new arguments are needed here. Our strategy towards hypercontractivity is detailed in Sections 3.2 and 4, and it actually represents the main task of the paper.

3.1. General strategy.

For the rest of the section, we fix two sequences $(X_i), (Y_i)$ of random variables in a free tracial probability space $(\mathcal{A},\varphi)$, two integers $N, m \ge 1$, as well as a function $f_N : \{1,\dots,N\}^d \to \mathbb{R}$ giving rise to a polynomial Q_N through (1), and we assume that all of these objects meet the requirements of Theorem 1.3. In accordance with the Lindeberg method, we are first prompted to introduce some additional notation.

Notation. For every $i \in \{1,\dots,N+1\}$, let us consider the vector

$$Z^{N,(i)} := (Y_1,\dots,Y_{i-1},X_i,\dots,X_N).$$

In particular, $Z^{N,(1)} = (X_1,\dots,X_N)$ and $Z^{N+1,(N)} = (Y_1,\dots,Y_N)$, so that

$$Q_N(X_1,\dots,X_N)^m - Q_N(Y_1,\dots,Y_N)^m = \sum_{i=1}^{N} \Big[Q_N(Z^{N,(i)})^m - Q_N(Z^{N,(i+1)})^m \Big]. \tag{22}$$

Since the only difference between the vectors $Z^{N,(i)}$ and $Z^{N,(i+1)}$ is their i^{th}-component, it is readily checked that

$$Q_N(Z^{N,(i)}) = U_N^{(i)} + V_N^{(i)}(X_i) \quad \text{and} \quad Q_N(Z^{N,(i+1)}) = U_N^{(i)} + V_N^{(i)}(Y_i),$$

where $U_N^{(i)}$ stands for the multilinear polynomial

$$U_N^{(i)} := \sum_{j_1,\dots,j_d \in \{1,\dots,N\}\backslash\{i\}} f_N(j_1,\dots,j_d)\, Z_{j_1}^{N,(i)} \dots Z_{j_d}^{N,(i)},$$

and $V_N^{(i)} : \mathcal{A} \to \mathcal{A}$ is the linear operator defined, for every $x \in \mathcal{A}$, by

$$V_N^{(i)}(x) := \sum_{l=1}^{d} \sum_{j_1,\dots,j_{d-1} \in \{1,\dots,N\}\backslash\{i\}} f_N(j_1,\dots,j_{l-1},i,j_l\dots,j_{d-1})\, Z_{j_1}^{N,(i)} \dots Z_{j_{l-1}}^{N,(i)} x Z_{j_l}^{N,(i)} \dots Z_{j_{d-1}}^{N,(i)}.$$

Expansion. Once endowed with the above notation, the problem reduces to examining the differences

$$\varphi\big((U_N^{(i)} + V_N^{(i)}(X_i))^m\big) - \varphi\big((U_N^{(i)} + V_N^{(i)}(Y_i))^m\big) \tag{23}$$

for $i \in \{1,\dots,N-1\}$. In a commutative context, this could be handled with the classical binomial formula. Although such a mere formula is not available here, one can still assert that for every $A, B \in \mathcal{A}$,

$$(A+B)^m = A^m + \sum_{n=1}^{m} \sum_{(r,\mathbf{i}_{r+1},\mathbf{j}_r) \in \mathcal{D}_{m,n}} c_{m,n,r,\mathbf{i}_{r+1},\mathbf{j}_r}\, A^{i_1} B^{j_1} A^{i_2} B^{j_2} \dots A^{i_r} B^{j_r} A^{i_{r+1}}, \tag{24}$$

where

$$\mathcal{D}_{m,n} := \{(r,\mathbf{i}_{r+1},\mathbf{j}_r) \in \{1,\dots,m\} \times \mathbb{N}^{r+1} \times \mathbb{N}^r : \sum_{l=1}^{r} j_l = n\ ,\ \sum_{l=1}^{r+1} i_l = m-n\}$$

and the $c_{m,n,r,\mathbf{i}_{r+1},\mathbf{j}_r}$'s stand for appropriate combinatorial coefficients (independent on A and B). The sets $\mathcal{D}_{m,n}$ must of course be understood as follows: given $(r,\mathbf{i}_{r+1},\mathbf{j}_r) \in \mathcal{D}_{m,n}$, the product $A^{i_1} B^{j_1} A^{i_2} B^{j_2} \dots A^{i_r} B^{j_r} A^{i_{r+1}}$ contains A exactly n times and B exactly $(m-n)$ times, both counted with multiplicity.

Let us go back to (23) and let us apply Formula (24) in order to expand $(U_N^{(i)} + V_N^{(i)}(X_i))^m$ (resp. $(U_N^{(i)} + V_N^{(i)}(Y_i))^m$). The first and second order terms (i.e., for $n = 1,2$ in (24)) of the resulting sum happen to vanish, as a straightforward use of the following lemma shows.

Lemma 3.1. *Let Y and Z be two centered random variables with unit variance. Then, for every integer $k \geq 1$ and every sequence (X_i) of centered freely independent random variables independent of Y and Z, one has*

$$\varphi\left(X_{i_1} \dots X_{i_r} Y X_{i_{r+1}} \dots X_{i_k}\right) = \varphi\left(X_{i_1} \dots X_{i_r} Z X_{i_{r+1}} \dots X_{i_k}\right) = 0 \tag{25}$$

and

$$\varphi\left(X_{i_1} \dots X_{i_r} Y X_{i_{r+1}} \dots X_{i_s} Y X_{i_{s+1}} \dots X_{i_k}\right) = \varphi\left(X_{i_1} \dots X_{i_r} Z X_{i_{r+1}} \dots X_{i_s} Z X_{i_{s+1}} \dots X_{i_k}\right) \tag{26}$$

for all $0 \leq r \leq s \leq k$ and $(i_1,\dots,i_k) \in \mathbb{N}^k$.

Proof. Let us first focus on (25). For $k = 1$, this is obvious. Assume that the result holds true up to $k-1$ and write

$$\varphi\big(X_{i_1} \dots X_{i_r} Y X_{i_{r+1}} \dots X_{i_k}\big) = \varphi\big(X_{i'_1}^{m_1} \dots X_{i'_{r'}}^{m_{r'}} Y X_{i'_{r'+1}}^{m_{r'+1}} \dots X_{i'_{s'}}^{m_{s'}}\big)$$

with $i'_{p+1} \neq i'_p$ for $p \in \{1, \dots, s'-1\} \backslash \{r'\}$, $i'_{s'} \neq i'_1$ and $m_p \geq 1$ for every $p \in \{1, \dots, s'\}$. Center successively every random variable $X_{i'_{p_1}}^{m_{p_1}}, \dots, X_{i'_{p_t}}^{m_{p_t}}$ for which $m_{p_i} \geq 2$: together with an induction argument, this yields

$$\begin{aligned}
&\varphi\big(X_{i'_1}^{m_1} \dots X_{i'_{r'}}^{m_{r'}} Y X_{i'_{r'+1}}^{m_{r'+1}} \dots X_{i'_{s'}}^{m_{s'}}\big) \\
&\quad = \varphi\big(X_{i'_1} \dots X_{i'_{p_1-1}} \big(X_{i'_{p_1}}^{m_{p_1}} - \varphi(X_{i'_{p_1}}^{m_{p_1}})\big) X_{i'_{p_1+1}}^{m_{p_1+1}} \dots X_{i'_{r'}}^{m_{r'}} Y X_{i'_{r'+1}}^{m_{r'+1}} \dots X_{i'_{s'}}^{m_{s'}}\big) \\
&\quad = \varphi\big(X_{i'_1} \dots X_{i'_{p_1-1}} \big(X_{i'_{p_1}}^{m_{p_1}} - \varphi(X_{i'_{p_1}}^{m_{p_1}})\big) X_{i'_{p_1+1}} \dots X_{i'_{p_2-1}} \big(X_{i'_{p_2}}^{m_{p_2}} - \varphi(X_{i'_{p_2}}^{m_{p_2}})\big) \\
&\qquad\qquad X_{i'_{p_2+1}}^{m_{p_2+1}} \dots X_{i'_{r'}}^{m_{r'}} Y X_{i'_{r'+1}}^{m_{r'+1}} \dots X_{i'_{s'}}^{m_{s'}}\big) = \dots = 0
\end{aligned}$$

owing to free independence. Identity (26) can be easily derived from a similar induction procedure. $\square$

Let us go back to the proof of Theorem 1.3. As a consequence of the previous lemma, it now suffices to establish that, either for $W = X_i$ or for $W = Y_i$, one has, as soon as $\sum_l j_l \geq 3$,

$$|\varphi\big((U_N^{(i)})^{i_1} (V_N^{(i)}(W))^{j_1} (U_N^{(i)})^{i_2} (V_N^{(i)}(W))^{j_2} \dots (U_N^{(i)})^{i_r} (V_N^{(i)}(W))^{j_r}\big)| \leq c_{m,d} \operatorname{Inf}_i(f_N)^{3/2} \tag{27}$$

for some constant $c_{m,d}$. Indeed, in this case, by combining (22), (24) and (27) with the identities in the statement of Lemma 3.1, we get

$$\begin{aligned}
|\varphi\big(Q_N(X_1, \dots, X_N)^m\big) - \varphi\big(Q_N(Y_1, \dots, Y_N)^m\big)| &\leq C_{m,d} \sum_{i=1}^N \operatorname{Inf}_i(f_N)^{3/2} \\
&\leq C_{m,d}\, \tau_N^{1/2} \sum_{i=1}^N \operatorname{Inf}_i(f_N) = C_{m,d}\, \tau_N^{1/2},
\end{aligned}$$

which is precisely the expected bound of Theorem 1.3.
In order to prove (27), let us first resort to the following Hölder-type inequality, borrowed from [3, Lemma 12]:

$$\begin{aligned}
&|\varphi\big((U_N^{(i)})^{i_1} (V_N^{(i)}(W))^{j_1} \dots (U_N^{(i)})^{i_r} (V_N^{(i)}(W))^{j_r}\big)| \\
&\quad \leq \varphi\big((U_N^{(i)})^{2^r i_1}\big)^{2^{-r}} \varphi\big((V_N^{(i)}(W))^{2^r j_1}\big)^{2^{-r}} \dots \varphi\big((U_N^{(i)})^{2^r i_r}\big)^{2^{-r}} \varphi\big((V_N^{(i)}(W))^{2^r j_r}\big)^{2^{-r}}.
\end{aligned} \tag{28}$$

Now, let the key (forthcoming) Proposition 3.3 come into the picture. Thanks to it, we can simultaneously assert that, for every $p \geq 1$,

$$\varphi\big((U_N^{(i)})^{2p}\big) \leq C_{p,d} \quad \text{and} \quad \varphi\big(V_N^{(i)}(X_i)^{2p}\big) \leq C_{p,d} \cdot \operatorname{Inf}_i(f_N)^p,$$

for some constant $C_{p,d}$. Going back to (28), we deduce that for every (j_l) such that $\sum_l j_l \geq 3$,

$$\begin{aligned}
|\varphi\big((U_N^{(i)})^{i_1} (V_N^{(i)}(X_i))^{j_1} \dots (U_N^{(i)})^{i_r} (V_N^{(i)}(X_i))^{j_r}\big)| &\leq C'_{r,d} \cdot \operatorname{Inf}_i(f_N)^{2^{-1}(j_1+\dots+j_r)} \\
&\leq C'_{r,d} \cdot \operatorname{Inf}_i(f_N)^{3/2}
\end{aligned}$$

since $\operatorname{Inf}_i(f_N) \leq 1$, and so the proof of Theorem 1.3 is done.

3.2. **Hypercontractivity.** Let us turn to the proof of the cornerstone result in the above strategy, namely the hypercontractivity property for homogeneous sums of free random variables. To be more specific, we shall show here how this property can be derived from the technical results contained in Section 4. The following elementary lemma will also play a role at some point in the sequel.

Lemma 3.2. *For every integer $r \geq 1$ and every sequence $X = (X_i)$ of random variables, one has $|\varphi\big(X_{i_1} \dots X_{i_{2r}}\big)| \leq \mu^X_{2^{r-1}}$, where $\mu^X_k := \sup_{1 \leq l \leq k\,,\, i \geq 1} \varphi\big(X_i^{2l}\big)$.*

Proof. For $r=1$, this corresponds to Cauchy-Schwarz inequality (see [8]). Assume that the result holds true up to $r-1$ ($r\geq 2$) for any sequence of random variables. By using Cauchy-Schwarz inequality, we first get

$$\begin{aligned}
&|\varphi\big(X_{i_1}\dots X_{i_{2r}}\big)|\\
&\quad= \ |\varphi\big((X_{i_1}\dots X_{i_r})(X_{i_{r+1}}\dots X_{i_{2r}})\big)|\\
&\quad\leq \ \varphi\big(X_{i_1}^2\dots X_{i_{r-1}}X_{i_r}^2X_{i_{r-1}}\dots X_{i_2}\big)^{1/2}\varphi\big(X_{i_{r+1}}^2\dots X_{i_{2r-1}}X_{i_{2r}}^2X_{i_{2r-1}}\dots X_{i_{r+2}}\big)^{1/2}. \qquad (29)
\end{aligned}$$

Denote by X^2 the sequence $X_1, X_1^2, X_2, X_2^2, \dots$. Then by induction, we deduce from (29) that $|\varphi\big(X_{i_1}\dots X_{i_{2r}}\big)|\leq \mu_{2^r-2}^{X^2}\leq \mu_{2^r-1}^{X}$, which achieves the proof. □

Proposition 3.3. *Let $X_1,\dots,X_N$ be centered freely independent random variables and denote by (μ_k^N) the sequence of larger even moments, i.e., $\mu_k^N := \sup_{1\leq i\leq N, 1\leq l\leq k}\varphi\big(X_i^{2l}\big)$. Fix $d\geq 1$, and consider a sequence of functions $f_N:\{1,\dots,N\}^d\to\mathbb{R}$ satisfying the three basic assumptions (i)-(ii)-(iii) of Theorem 1.3. Define Q_N through (1). Then for every $r\geq 1$, there exists a constant $C_{r,d}$ such that*

$$\varphi\big(Q_N(X_1,\dots,X_N)^{2r}\big)\leq C_{r,d}\,\mu_{2^{rd-1}}^N\Bigg(\sum_{j_1,\dots,j_d=1}^N f_N(j_1,\dots,j_d)^2\Bigg)^r. \qquad (30)$$

Proof. Owing to Lemma 3.1, it holds that

$$\begin{aligned}
&\varphi\big(Q_N(X_1,\dots,X_N)^{2r}\big)\\
&\quad= \sum_{\substack{1\leq j_1^1,\dots,j_d^1\leq N\\ \vdots\\ 1\leq j_1^{2r},\dots,j_d^{2r}\leq N}} f_N(j_1^1,\dots,j_d^1)\dots f_N(j_1^{2r},\dots,j_d^{2r})\varphi\big((X_{j_1^1}\dots X_{j_d^1})\dots(X_{j_1^{2r}}\dots X_{j_d^{2r}})\big)\\
&\quad= \sum_{(j_1^1,\dots,j_d^{2r})\in\mathcal{A}_{2rd}^N} f_N(j_1^1,\dots,j_d^1)\dots f_N(j_1^{2r},\dots,j_d^{2r})\varphi\big((X_{j_1^1}\dots X_{j_d^1})\dots(X_{j_1^{2r}}\dots X_{j_d^{2r}})\big),
\end{aligned}$$

where we have set, for every $R\geq 1$,

$$\mathcal{A}_R^N := \{(j_1,\dots,j_R)\in\{1,\dots,N\}^R:\ \text{for each } i_1,\ \text{there exists } i_2\neq i_1 \text{ such that } j_{i_1}=j_{i_2}\}.$$

Bound each term of the form $\varphi\big((X_{j_1^1}\dots X_{j_d^1})\dots(X_{j_1^{2r}}\dots X_{j_d^{2r}})\big)$ of this sum by means of Lemma 3.2 to get

$$\varphi\big(Q_N(X_1,\dots,X_N)^{2r}\big)\leq \mu_{2^{rd-1}}^N\sum_{(j_1^1,\dots,j_d^{2r})\in\mathcal{A}_{2rd}^N}|f_N(j_1^1,\dots,j_d^1)|\dots|f_N(j_1^{2r},\dots,j_d^{2r})|.$$

Now write

$$\begin{aligned}
&\sum_{(j_1^1,\dots,j_d^{2r})\in\mathcal{A}_{2rd}^N}|f_N(j_1^1,\dots,j_d^1)|\dots|f_N(j_1^{2r},\dots,j_d^{2r})|\\
&\qquad\qquad= \sum_{d\leq k\leq rd}\ \sum_{(j_1^1,\dots,j_d^{2r})\in\mathcal{A}_{2rd}^{N,k}}|f_N(j_1^1,\dots,j_d^1)|\dots|f_N(j_1^{2r},\dots,j_d^{2r})|,
\end{aligned}$$

where $\mathcal{A}_R^{N,k} := \{(j_1,\dots,j_R)\in\mathcal{A}_R^N:\ |\{j_1,\dots,j_R\}|=k\}$. For each $(j_1,\dots,j_d^{2r})\in\mathcal{A}_{2rd}^{N,k}$, one has $\{j_1^1,\dots,j_d^{2r}\}=\{j_1',\dots,j_k'\}$ and

$$|f_N(j_1^1,\dots,j_d^1)|\dots|f_N(j_1^{2r},\dots,j_d^{2r})| = \prod_{(i_1,\dots,i_d)\in\{1,\dots,k\}^d}|f(j_{i_1}',\dots,j_{i_d}')|^{m_{i_1,\dots,i_d}},$$

for some weights $m_{i_1,\dots,i_d} \geq 0$ that satisfy the three following conditions:
(i) $\sum_{(i_1,\dots,i_d)\in\{1,\dots,k\}^d} m_{i_1,\dots,i_d} = 2r$;
(ii) $m_{i_1,\dots,i_d} = 0$ if $i_l = i_k$ for some $l \neq k$ (since f_N vanishes on diagonals);
(iii) $\sum_{\substack{i_1,\dots,i_d\in\{1,\dots,k\}\\ i\in\{i_1,\dots,i_d\}}} m_{i_1,\dots,i_d} \geq 2$ for all i (since $(j_1^1,\dots,j_d^{2r}) \in \mathcal{A}^N_{2rd}$).
As a result,

$$\sum_{(j_1^1,\dots,j_d^{2r})\in\mathcal{A}^{N,k}_{2rd}} |f_N(j_1^1,\dots,j_d^1)|\dots|f_N(j_1^{2r},\dots,j_d^{2r})|$$

$$\leq \; c_{r,d} \sum_{d\leq k\leq rd} \;\sum_{(m_{i_1,\dots,i_d})\in\mathcal{M}^{2r}_{d,k}} \left(\sum_{j_1,\dots,j_k=1}^{N} \prod_{(i_1,\dots,i_d)\in\{1,\dots,k\}^d} |f_N(j_{i_1},\dots,j_{i_d})|^{m_{i_1,\dots,i_d}} \right),$$

where $\mathcal{M}^{2r}_{d,k} := \{(m_{i_1,\dots,i_d})_{1\leq i_1,\dots,i_d\leq k} \in \mathbb{N}^{kd}$ for which (i)-(ii)-(iii) holds true$\}$ and $c_{r,d}$ is a constant that only depends on r and d. At this point, we are in a position to apply the forthcoming Corollary 4.2 (for fixed k and $(m_{i_1,\dots,i_d}) \in \mathcal{M}^{2r}_{d,k}$) to assert that

$$\sum_{j_1,\dots,j_k=1}^{N} \prod_{(i_1,\dots,i_d)\in\{1,\dots,k\}^d} |f_N(j_{i_1},\dots,j_{i_d})|^{m_{i_1,\dots,i_d}} \leq \left(\sum_{j_1,\dots,j_d=1}^{N} f_N(j_1,\dots,j_d)^2 \right)^r \tag{31}$$

and this achieves the proof of (30). $\square$

4. A technical result towards hypercontractivity

We are now left with the proof of (31). This estimate will actually be derived from a more general inequality. In order to state it, we introduce some further notation.

Consider two integers $k \geq d \geq 1$ together with a sequence $(m_{i_1,\dots,i_d})_{1\leq i_1,\dots,i_d\leq k}$ of positive weights vanishing on diagonals. Let us also fix a k-uple of indices $(j_1,\dots,j_k) \in \{1,\dots,N\}^k$, as well as a function $f : \{1,\dots,N\}^d \to \mathbb{R}$. If $J_1,\dots,J_s$ ($s \leq d$) are non-empty disjoint subsets of $\{j_1,\dots,j_k\}$, we denote by

$$f(J_1,\dots,J_s)$$

the product of all of the terms $f(j_{i_1},\dots,j_{i_d})^{m_{i_1,\dots,i_d}}$ that only appeal to $j_{i_1},\dots,j_{i_d} \in J_1\cup\dots\cup J_s$, with at least one j_{i_l} in each J_i ($i = 1,\dots,s$).
For instance, if $d = 2$, $J_1 := \{j_1, j_2\}$, $J_2 := \{j_3, j_4\}$, one has

$$f(J_1) = f(j_1,j_2)^{m_{1,2}} f(j_2,j_1)^{m_{2,1}} \quad , \quad f(J_1,J_2) = \prod_{\substack{i_1\in\{1,2\}\\ i_2\in\{3,4\}}} f(j_{i_1},j_{i_2})^{m_{i_1,i_2}} f(j_{i_2},j_{i_1})^{m_{i_2,i_1}}. \tag{32}$$

For the sake of consistency, we also set $f(J_1,\dots,J_s) = 1$ in each of the following situations: (i) $s > d$, (ii) one of the sets J_i is empty, (iii) the cardinal of $J_1\cup\dots\cup J_s$ is strictly smaller than d.
With this notation in hand, observe that when splitting a block J_1 into disjoint (non-empty) subsets $J_{1,1}, J_{1,2}$, the product $f(J_1,\dots,J_s)$ must be decomposed as

$$f(J_1,\dots,J_s) = f(J_{1,1},J_2,\dots,J_s) f(J_{1,2},J_2,\dots,J_s) f(J_{1,1},J_{1,2},J_2,\dots,J_s),$$

and similar formulas hold for divisions of J_1 into k blocks ($k \geq 2$). Such a splitting will be extensively used in the proof of Proposition 4.1.

Besides, we denote by $m_{J_1,\ldots,J_s}$ the sum of the weights involved in $f(J_1,\ldots,J_s)$. To take up with the above example (32), one has of course

$$m_{J_1} = m_{1,2} + m_{2,1} \quad , \quad m_{J_1,J_2} = \sum_{\substack{i_1\in\{1,2\}\\ i_2\in\{3,4\}}} \{m_{i_1,i_2} + m_{i_2,i_1}\}\, .$$

As far as summands are concerned, we will often use the convenient *block* convention $\sum_{\{j_1,\ldots,j_k\}}$ for $\sum_{j_1,\ldots,j_k=1}^N$. Finally, the notation $|J_i|$ will refer to the number of indices contained in J_i.

We are now in a position to state the main result of this section:

Proposition 4.1. *Assume that $d \geq 1$, $0 \leq l \leq k$ and $k \geq d$. Let $(m_{i_1,\ldots,i_d})_{1\leq i_1,\ldots,i_d\leq k}$ be positive weights vanishing on the diagonals such that*

$$\sum_{\substack{1\leq i_1,\ldots,i_d\leq k\\ i\in(i_1,\ldots,i_d)}} m_{i_1,\ldots,i_d} \geq 1 \quad \text{if} \quad 1\leq i\leq l \quad \text{and} \quad \sum_{\substack{1\leq i_1,\ldots,i_d\leq k\\ i\in(i_1,\ldots,i_d)}} m_{i_1,\ldots,i_d} \geq 2 \quad \text{if} \quad l+1\leq i\leq k. \tag{33}$$

Then, with the above notational convention, one has, for every function $f : \{1,\ldots,N\}^d \to \mathbb{R}$,

$$\sum_{j_1,\ldots,j_l=1}^N f(\{j_1,\ldots,j_l\})^2 \Bigg(\sum_{j_{l+1},\ldots,j_k=1}^N f(\{j_1,\ldots,j_l\},\{j_{l+1},\ldots,j_k\}) f(\{j_{l+1},\ldots,j_k\})\Bigg)^2$$
$$\leq \Bigg(\sum_{i_1,\ldots,i_d=1}^N f(i_1,\ldots,i_d)^2\Bigg)^{\sum_{1\leq i_1,\ldots,i_d\leq k} m_{i_1,\ldots,i_d}}. \tag{34}$$

As an immediate spin-off of this result, we deduce the following estimate (take $l = 0$ in (34)), which leads to (31) in a obvious way.

Corollary 4.2. *Assume that $d \geq 1$, $k \geq d$, and let $(m_{i_1,\ldots,i_d})_{1\leq i_1,\ldots,i_d\leq k}$ be positive weights vanishing on the diagonals such that*

$$\sum_{\substack{1\leq i_1,\ldots,i_d\leq k\\ i\in(i_1,\ldots,i_d)}} m_{i_1,\ldots,i_d} \geq 2 \quad \text{for every} \quad 1\leq i\leq k.$$

Then for every function $f : \{1,\ldots,N\}^d \to \mathbb{R}$, one has

$$\Bigg(\sum_{j_1,\ldots,j_k=1}^N \prod_{i_1,\ldots,i_d\in\{1,\ldots,k\}} f(j_{i_1},\ldots,j_{i_d})^{m_{i_1,\ldots,i_d}}\Bigg)^2 \leq \Bigg(\sum_{j_1,\ldots,j_d=1}^N f(j_1,\ldots,j_d)^2\Bigg)^{\sum_{1\leq i_1,\ldots,i_d\leq k} m_{i_1,\ldots,i_d}}. \tag{35}$$

Remark 4.3. The condition (33) on the weights must be understood as follows in the left-hand-side of (34). On the one hand, every index j_i from the first summand (i.e., $1 \leq i \leq l$) comes out at least once in $f(\{j_1,\ldots,j_l\})$ or in $f(\{j_1,\ldots,j_l\},\{j_{l+1},\ldots,j_k\})$ (it may also appear in both expressions). On the other hand, every index j_i from the second summand (i.e., $l+1 \leq i \leq k$) comes out at least twice in

$$f(\{j_1,\ldots,j_l\},\{j_{l+1},\ldots,j_k\}) f(\{j_{l+1},\ldots,j_k\}),$$

taking into account the multiplicativity induced by the weights (see the subsequent example).

Remark 4.4. At first sight, the reader might (legitimately) wonder why we focus on the general estimate (34) for the purpose of the paper. Indeed, according to the proof of Proposition 3.3, we know that the bound (35) suffices to achieve the strategy towards Theorem 1.3, and (35) is only

a very particular case (much easier to handle *a priori*) of (34). In fact, as we tried to estimate the left-hand-side of (35), it soon appeared to us that the more general structure

$$\sum_{J_1} f(J_1)^2 \Big(\sum_{J_2} f(J_1, J_2) f(J_2) \Big)^2 \tag{36}$$

involved in (34) should play the role of a *recurring pattern* throughout the procedure. This idea should become clear in the course of the proof, especially by considering the so-called Config. 8 in the fully-symmetric situation (see also the following example).

The basic ingredient towards (34) essentially lies in a blockwise application of the Cauchy-Schwarz inequality (together with some even more trivial estimates). To get an idea on how things should work, consider the simple example

$$\sum_{j_1,j_2,j_3} f(j_1,j_2)^2 \Big(\sum_{j_4,j_5,j_6,j_7} f(j_2,j_4) f(j_2,j_5) f(j_3,j_4) f(j_3,j_6) f(j_5,j_6) f(j_6,j_7)^2 \Big)^2 . \tag{37}$$

This expression indeed fits the pattern of the left-hand-side in (34) (see Remark 4.3): for $1 \leq i \leq 3$, j_i appears at least once, while j_4, j_5, j_6, j_7 all appear at least twice (note that j_7 is counted twice in $f(j_6, j_7)^2$ according to our convention). Then, by Cauchy-Schwarz inequality (over j_5, j_6), we get

$$\begin{aligned}
&\sum_{j_1,j_2,j_3} f(j_1,j_2)^2 \Big(\sum_{j_4,j_5,j_6,j_7} f(j_2,j_4) f(j_2,j_5) f(j_3,j_4) f(j_3,j_6) f(j_5,j_6) f(j_6,j_7)^2 \Big)^2 \\
&= \sum_{j_1,j_2,j_3} f(j_1,j_2)^2 \Big(\sum_{j_5,j_6} \Big[\sum_{j_4} f(j_2,j_4) f(j_2,j_5) f(j_3,j_4) f(j_3,j_6) \Big] \Big[f(j_5,j_6) \sum_{j_7} f(j_6,j_7)^2 \Big] \Big)^2 \\
&\leq \Big[\sum_{j_1,j_2,j_3,j_5,j_6} f(j_1,j_2)^2 \Big(\sum_{j_4} f(j_2,j_4) f(j_2,j_5) f(j_3,j_4) f(j_3,j_6) \Big)^2 \Big] \times \\
&\qquad\qquad \Big[\sum_{j_5,j_6} f(j_5,j_6)^2 \Big(\sum_{j_7} f(j_5,j_7)^2 \Big)^2 \Big].
\end{aligned}$$

At this point, one realizes that the brackets in the latter bound both fit the pattern of (36), too. Besides, in comparison to (37), fewer indices j_i now come into the picture: the first bracket appeals to 6 indices (namely $j_1, j_2, j_3, j_5, j_6, j_4$), while the second bracket only appeal to j_5, j_6, j_7. This stability phenomenon allows us to settle an iteration procedure on the number k of indices j_i involved in (36), and the problem accordingly reduces to the initialization step, i.e., $k = d$. Unfortunately, when implementing this elementary strategy in a general setting, one has to cope with possible "degenerate" configurations of (36), for which Cauchy-Schwarz inequality cannot be readily applied as such. This makes the proof of Proposition 4.1 quite long and technical, although it only appeals to very basic estimates.

For the sake of clarity, we have divided this proof into two steps. First, we deal with the particular case where f is a fully-symmetric function. Then we tackle the general situation, for which only slight technical modifications are required.

Proof of Proposition 4.1 in the fully-symmetric case. We assume here that the functions f involved in the proof are all fully-symmetric.

Consider first an induction procedure on the number d of arguments for f. When $d=1$, the result is obvious. Indeed, one has in this case

$$\begin{aligned}
&\sum_{j_1,\ldots,j_l} f(j_1)^{2m_1}\ldots f(j_l)^{2m_l}\Big(\sum_{j_{l+1},\ldots,j_k} f(j_{l+1})^{m_{l+1}}\ldots f(j_k)^{m_k}\Big)^2\\
&= \Big(\sum_{j_1} f(j_1)^{2m_1}\Big)\ldots\Big(\sum_{j_l} f(j_l)^{2m_l}\Big)\Big(\sum_{j_{l+1}} f(j_{l+1})^{m_{l+1}}\Big)^2\ldots\Big(\sum_{j_k} f(j_k)^{m_k}\Big)^2\\
&\le \Big(\sum_i f(i)^2\Big)^{m_1+\ldots+m_k}
\end{aligned}$$

since $m_i \ge 1$ for $i\in\{1,\ldots,l\}$ and $m_i\ge 2$ for $i\in\{l+1\ldots,k\}$.

From now on, we assume that the estimate (34) holds true for every function with less than $d-1$ arguments ($d\ge 2$) and we consider $f:\{1,\ldots,N\}^d\to\mathbb{R}$. To extend the result to f, we turn to a second induction procedure on the number k of indices j_i involved in (34).

Initialization step: $k=d$. By setting $m_{[1,\ldots,d]} := \sum_{\sigma\in\mathfrak{S}_d} m_{\sigma(1),\ldots,\sigma(d)}$, the only possible configurations are

$$\sum_{j_1,\ldots,j_l}\Big(\sum_{j_{l+1},\ldots,j_d} f(j_1,\ldots,j_d)^{m_{[1,\ldots,d]}}\Big)^2,\quad \text{where } 0\le l\le d-1 \text{ and } m_{[1,\ldots,d]}\ge 2, \tag{38}$$

and

$$\sum_{j_1,\ldots,j_d} f(j_1,\ldots,j_d)^{2m_{[1,\ldots,d]}}\quad \text{with } m_{[1,\ldots,d]}\ge 1. \tag{39}$$

As far as (38) is concerned, one has, since $m_{[1,\ldots,d]}\ge 2$,

$$\begin{aligned}
\sum_{j_1,\ldots,j_l}\Big(\sum_{j_{l+1},\ldots,j_d} f(j_1,\ldots,j_d)^{m_{[1,\ldots,d]}}\Big)^2 &\le \sum_{j_1,\ldots,j_l}\Big(\sum_{j_{l+1},\ldots,j_d} \big|f(j_1,\ldots,j_d)\big|^{m_{[1,\ldots,d]}}\Big)^2\\
&\le \Big(\sum_{j_1,\ldots,j_d} \big|f(j_1,\ldots,j_d)\big|^{m_{[1,\ldots,d]}}\Big)^2\\
&\le \Big(\sum_{j_1,\ldots,j_d} f(j_1,\ldots,j_d)^2\Big)^{m_{[1,\ldots,d]}},
\end{aligned}$$

which is the expected bound. The estimate of (39) is obvious.

From now on, we assume that the result holds true up to $k-1$ indices ($k\ge d+1$). To extend it to the k-index situation, we are prompted to successively examine different well-chosen configurations for the left-hand side of (34): these are the subsequent Config. 1-8. Observe *a posteriori* that some of these situations are included in others. In fact, it is readily checked that the two configurations 6 and 8 cover (together) all of the possible combinations for (36), which ensures the achievement of our procedure.

Config. 1: $\sum_{j_1,J_1} f(\{j_1\},J_1)^2\Big(\sum_{J_2} f(\{j_1\},J_1,J_2)f(\{j_1\},J_2)\Big)^2$, with $|J_1|+|J_2|=k-1$. In other words, we suppose here that the index j_1 appears in each term. As f is assumed to be fully-symmetric, we can write $f(j_1,J_1)$ for $f(\{j_1\},J_1)$ to indicate that j_1 can always be put in the first position. Then, by applying the induction hypothesis (on d) to the function

$(i_2,\dots,i_d)\mapsto f(j_1,i_2,\dots,i_d)$, we deduce that

$$\begin{aligned}
&\sum_{j_1,J_1} f(j_1,J_1)^2\Big(\sum_{J_2} f(j_1,J_1,J_2)f(j_1,J_2)\Big)^2\\
&\quad= \sum_{j_1}\Big[\sum_{J_1} f(j_1,J_1)^2\Big(\sum_{J_2} f(j_1,J_2)f(j_1,J_1,J_2)\Big)^2\Big]\\
&\quad\le \sum_{j_1}\Big[\sum_{i_2,\dots,i_d} f(j_1,i_2,\dots,i_d)^2\Big]^{m_{\{j_1\},J_1}+m_{\{j_1\},J_1,J_2}+m_{\{j_1\},J_2}}\\
&\quad\le \Big(\sum_{i_1,i_2,\dots,i_d} f(i_1,i_2,\dots,i_d)^2\Big)^{m_{\{j_1\},J_1}+m_{\{j_1\},J_1,J_2}+m_{\{j_1\},J_2}},
\end{aligned}$$

which clearly corresponds to the expected bound in this situation.

Config. 2: $\sum_{J_1} f(J_1)^2$ with $|J_1|=k$, i.e. $l=k$ in (34). Write first J_1 as the disjoint union $J_1=\{j_1\}\cup J_1'$ and accordingly

$$\sum_{J_1} f(J_1)^2=\sum_{j_1,J_1'} f(\{j_1\},J_1')^2 f(J_1')^2.$$

Then we divide J_1' into the disjoint union $J_1'=J_{1,1}\cup J_{1,2}\cup J_{1,3}$, where

$$J_{1,1}:=\{j_i\in J_1' \text{ that appears in } f(\{j_1\},J_1')^2 \text{ but not in } f(J_1')^2\},$$

$$J_{1,2}:=\{j_i\in J_1' \text{ that appears both in } f(\{j_1\},J_1')^2 \text{ and in } f(J_1')^2\},$$

$$J_{1,3}:=\{j_i\in J_1' \text{ that appears in } f(J_1')^2 \text{ but not in } f(\{j_1\},J_1')^2\}.$$

In this way, we get

$$\begin{aligned}
&\sum_{j_1,J_1'} f(\{j_1\},J_1')^2 f(J_1')^2\\
&\quad= \sum_{J_{1,2}}\Big[\sum_{j_1,J_{1,1}} f(\{j_1\},J_{1,1})^2 f(\{j_1\},J_{1,2})^2 f(\{j_1\},J_{1,1},J_{1,2})^2\Big]\Big[f(J_{1,2})^2\sum_{J_{1,3}} f(J_{1,2},J_{1,3})^2 f(J_{1,3})^2\Big]\\
&\quad\le \Big[\sum_{j_1,J_{1,1},J_{1,2}} f(\{j_1\},J_{1,1})^2 f(\{j_1\},J_{1,2})^2 f(\{j_1\},J_{1,1},J_{1,2})^2\Big]\times\\
&\qquad\qquad\Big[\sum_{J_{1,2},J_{1,3}} f(J_{1,2})^2 f(J_{1,2},J_{1,3})^2 f(J_{1,3})^2\Big],
\end{aligned}\tag{40}$$

where we have only used the trivial estimate $\sum_i a_ib_i\le(\sum_i a_i)(\sum_i b_i)$ when $a_i,b_i\ge 0$. The first bracket in (40) is a particular case of Config. 1 (take $J_2=\emptyset$), which yields

$$\begin{aligned}
&\sum_{j_1,J_{1,1},J_{1,2}} f(\{j_1\},J_{1,1})^2 f(\{j_1\},J_{1,2})^2 f(\{j_1\},J_{1,1},J_{1,2})^2\\
&\qquad\le \Big(\sum_{i_1,\dots,i_d} f(i_1,\dots,i_d)^2\Big)^{m_{\{j_1\},J_{1,1}}+m_{\{j_1\},J_{1,1},J_{1,2}}+m_{\{j_1\},J_{1,2}}}.
\end{aligned}\tag{41}$$

As for the second bracket, observe first that it fits the pattern of (36). Then, as $|J_{1,2}|+|J_{1,3}| \leq k-1$, we can apply the induction hypothesis on k to assert that

$$\sum_{J_{1,2},J_{1,3}} f(J_{1,2})^2 f(J_{1,2},J_{1,3})^2 f(J_{1,3})^2 \leq \Big(\sum_{i_1,\ldots,i_d} f(i_1,\ldots,i_d)^2\Big)^{m_{J_{1,2}}+m_{J_{1,2},J_{1,3}}+m_{J_{1,3}}}. \tag{42}$$

Going back to (40) and given the very definition of the subsets $J_{1,1}, J_{1,2}, J_{1,3}$, it is easy to realize that (41) and (42) provide us with the expected bound (observe for instance that $m_{J_{1,1},J_{1,3}}=0$, $m_{\{j_1\},J_{1,3}}=0$, etc).

Config. 3: $\sum_{j_1,J_1}\Big(\sum_{J_{2,1},J_{2,2}} f(\{j_1\},J_{2,1})f(\{j_1\},J_1,J_{2,1})f(J_1,J_{2,2})\Big)^2$ with $k=1+|J_1|+|J_{2,1}|+|J_{2,2}|$. This sum can also be written as

$$\sum_{J_1}\Big[\sum_{j_1}\Big(\sum_{J_{2,1}} f(\{j_1\},J_{2,1})f(\{j_1\},J_1,J_{2,1})\Big)^2\Big]\Big[\Big(\sum_{J_{2,2}} f(J_1,J_{2,2})\Big)^2\Big]. \tag{43}$$

Now split up J_1 into $J_1=J_{1,1}\cup J_{1,2}\cup J_{1,3}$ with

$$\begin{aligned}
J_{1,1} &:= \{j_i\in J_1 \text{ that appears in the first bracket of (43) but not in the second one}\},\\
J_{1,2} &:= \{j_i\in J_1 \text{ that appears in both brackets of (43)}\},\\
J_{1,3} &:= \{j_i\in J_1 \text{ that appears in the second bracket of (43) but not in the first one}\}.
\end{aligned}$$

With these subsets in hand, (43) clearly becomes

$$\begin{aligned}
&\sum_{J_{1,2}}\Big[\sum_{j_1,J_{1,1}}\Big(\sum_{J_{2,1}} f(\{j_1\},J_{2,1})f(\{j_1\},J_{1,1},J_{2,1})f(\{j_1\},J_{1,1},J_{1,2},J_{2,1})f(\{j_1\},J_{1,2},J_{2,1})\Big)^2\Big]\\
&\qquad\qquad\times\Big[\sum_{J_{1,3}}\Big(\sum_{J_{2,2}} f(J_{1,2},J_{2,2})f(J_{1,2},J_{1,3},J_{2,2})f(J_{1,3},J_{2,2})\Big)^2\Big]\\
&\qquad\qquad\qquad\qquad\leq\Big[\sum_{j_1,J_{1,1},J_{1,2}}\Big(\sum_{J_{2,1}}\cdots\Big)^2\Big]\Big[\sum_{J_{1,2},J_{1,3}}\Big(\sum_{J_{2,2}}\cdots\Big)^2\Big].
\end{aligned}\tag{44}$$

Here again, the first bracket in (44) is a particular case of Config. 1. As far as the second bracket is concerned, observe first that it fits the pattern of (36). Therefore, since $|J_{1,2}|+|J_{1,3}|+|J_{2,2}|\leq k-1$, we are allowed to use the induction hypothesis on k to bound it appropriately. As in Config. 2, the conclusion is now easily derived.

Config. 4: $\sum_{J_1}\Big(\sum_{J_2} f(J_1,J_2)\Big)^2$ with $J_1\neq\emptyset$ and $|J_1|+|J_2|=k$. If $J_1=\{j_1\}$, then the situation reduces to a particular case of Config. 1. Otherwise, write $J_1=\{j_1\}\cup J_{1,1}$ with $J_{1,1}\neq\emptyset$, so the sum becomes

$$\sum_{j_1,J_{1,1}}\Big(\sum_{J_2} f(\{j_1\},J_2)f(\{j_1\},J_{1,1},J_2)f(J_{1,1},J_2)\Big)^2.$$

Now, divide J_2 into $J_2=J_{2,1}\cup J_{2,2}\cup J_{2,3}$ with

$$\begin{aligned}
J_{2,1} &:= \{j_i\in J_2 \text{ that appears in } f(\{j_1\},J_2)f(\{j_1\},J_{1,1},J_2) \text{ but not in } f(J_{1,1},J_2)\},\\
J_{2,2} &:= \{j_i\in J_2 \text{ that appears both in } f(\{j_1\},J_2)f(\{j_1\},J_{1,1},J_2) \text{ and in } f(J_{1,1},J_2)\},\\
J_{2,3} &:= \{j_i\in J_2 \text{ that appears in } f(J_{1,1},J_2) \text{ but not in } f(\{j_1\},J_2)f(\{j_1\},J_{1,1},J_2)\}.
\end{aligned}$$

Thus, we have

$$\sum_{J_2} f(\{j_1\},J_2)f(\{j_1\},J_{1,1},J_2)f(J_{1,1},J_2) = \sum_{J_{2,2}}$$
$$\Big[f(\{j_1\},J_{2,2})f(\{j_1\},J_{1,1},J_{2,2})\sum_{J_{2,1}} f(\{j_1\},J_{2,1})f(\{j_1\},J_{2,1},J_{2,2})f(\{j_1\},J_{1,1},J_{2,1})f(\{j_1\},J_{1,1},J_{2,1},J_{2,2})\Big]$$
$$\times \Big[f(J_{1,1},J_{2,2})\sum_{J_{2,3}} f(J_{1,1},J_{2,2},J_{2,3})f(J_{1,1},J_{2,3})\Big].$$

If $J_{2,2}=\emptyset$, we are brought back to Config. 3. If $J_{2,2}\neq\emptyset$, then by Cauchy-Schwarz inequality (over $J_{2,2}$), we get

$$\sum_{j_1,J_{1,1}}\Big(\sum_{J_{2,2}}\Big[\cdots\Big]\Big[\cdots\Big]\Big)^2$$
$$\leq \sum_{J_{1,1}}\Big[\sum_{j_1,J_{2,2}} f(\{j_1\},J_{2,2})^2 f(\{j_1\},J_{1,1},J_{2,2})^2\Big(\sum_{J_{2,1}}\cdots\Big)^2\Big]\Big[\sum_{J_{2,2}} f(J_{1,1},J_{2,2})^2\Big(\sum_{J_{2,3}}\cdots\Big)^2\Big],$$

and from this bound, the conclusion is easily derived with the arguments of Config. 3 (we are in the same position as in (43)).

Config. 5: $\sum_{J_1} f(J_1)^2\Big(\sum_{J_2} f(J_1,J_2)\Big)^2$ with $J_1\neq\emptyset$ and $|J_1|+|J_2|=k$. Write $J_1=J_{1,1}\cup J_{1,2}\cup J_{1,3}$ with

$$\begin{aligned}
J_{1,1} &:= \{j_i\in J_1 \text{ that appears in } f(J_1)^2 \text{ but not in } f(J_1,J_2)\},\\
J_{1,2} &:= \{j_i\in J_1 \text{ that appears both in } f(J_1)^2 \text{ and in } f(J_1,J_2)\},\\
J_{1,3} &:= \{j_i\in J_1 \text{ that appears in } f(J_1,J_2) \text{ but not in } f(J_1)^2\}.
\end{aligned}$$

Then one has

$$\begin{aligned}
\sum_{J_1} f(J_1)^2\Big(\sum_{J_2} f(J_1,J_2)\Big)^2 &= \sum_{J_{1,2}}\Big[f(J_{1,2})^2\sum_{J_{1,1}} f(J_{1,1})^2 f(J_{1,1},J_{1,2})^2\Big]\times\\
&\quad \Big[\sum_{J_{1,3}}\Big(\sum_{J_2} f(J_{1,2},J_2)f(J_{1,2},J_{1,3},J_2)f(J_{1,3},J_2)\Big)^2\Big]\\
&\leq \Big[\sum_{J_{1,1},J_{1,2}} f(J_{1,1})^2 f(J_{1,1},J_{1,2})^2 f(J_{1,2})^2\Big]\times\\
&\quad \Big[\sum_{J_{1,2},J_{1,3}}\Big(\sum_{J_2} f(J_{1,2},J_2)f(J_{1,2},J_{1,3},J_2)f(J_{1,3},J_2)\Big)^2\Big].
\end{aligned}$$

The first bracket corresponds to Config. 2, while the second bracket is a particular case of Config. 4, which leads to the expected bound.

Config. 6: $\sum_{J_1} f(J_1)^2 \left(\sum_{J_2} f(J_1,J_2) f(J_2)\right)^2$ with $J_1 \neq \emptyset$ and $|J_1|+|J_2|=k$. This time, split up J_2 into $J_2 = J_{2,1} \cup J_{2,2} \cup J_{2,3}$ with

$$\begin{aligned} J_{2,1} &:= \{j_i \in J_2 \text{ that appears in } f(J_1,J_2) \text{ but not in } f(J_2)\}, \\ J_{2,2} &:= \{j_i \in J_2 \text{ that appears both in } f(J_1,J_2) \text{ and in } f(J_2)\}, \\ J_{2,3} &:= \{j_i \in J_2 \text{ that appears in } f(J_2) \text{ but not in } f(J_1,J_2)\}. \end{aligned}$$

Thus,

$$\sum_{J_2} f(J_1,J_2) f(J_2) = \\ \sum_{J_{2,2}} \left[f(J_1,J_{2,2}) \sum_{J_{2,1}} f(J_1,J_{2,1}) f(J_1,J_{2,1},J_{2,2}) \right] \left[f(J_{2,2}) \sum_{J_{2,3}} f(J_{2,2},J_{2,3}) f(J_{2,3}) \right].$$

If $J_{2,2} = \emptyset$, then we go back to Config. 5 for $\sum_{J_1} f(J_1)^2 \left(\sum_{J_{2,1}} f(J_1,J_{2,1}) \right)^2$ and we can use the induction hypothesis (on k) to cope with $\left(\sum_{J_{2,3}} f(J_{2,3}) \right)^2$ (since $|J_{2,3}| \leq |J_2| \leq k-1$). If $J_{2,2} \neq \emptyset$, one has by Cauchy-Schwarz inequality

$$\begin{aligned} &\sum_{J_1} f(J_1)^2 \left(\sum_{J_2} f(J_1,J_2) f(J_2) \right)^2 \\ &\leq \left[\sum_{J_1,J_{2,2}} f(J_1)^2 f(J_1,J_{2,2})^2 \left(\sum_{J_{2,1}} f(J_1,J_{2,1}) f(J_1,J_{2,1},J_{2,2}) \right)^2 \right] \\ &\qquad \times \left[\sum_{J_{2,2}} f(J_{2,2})^2 \left(\sum_{J_{2,3}} f(J_{2,2},J_{2,3}) f(J_{2,3}) \right)^2 \right]. \end{aligned}$$

Now, if $J_{2,3} \neq \emptyset$, we can rely on the induction hypothesis (on k) to conclude, whereas the situation reduces to Config. 2 and 5 if $J_{2,3} = \emptyset$.

Config. 7: $\left(\sum_{j_1,J_2} f(\{j_1\},J_2) \right)^2$ with $|J_2| = k-1$ (in particular, $l=0$ in (34)). Just as in Config. 1, we can here rely on the induction hypothesis on d to write

$$\begin{aligned} \left(\sum_{j_1,J_2} f(j_1,J_2) \right)^2 &\leq \left(\sum_{j_1} \left| \sum_{J_2} f(j_1,J_2) \right| \right)^2 \\ &\leq \left(\sum_{j_1} \left(\sum_{i_2,\ldots,i_d} f(j_1,i_2,\ldots,i_d)^2 \right)^{\frac{m_{\{j_1\},J_2}}{2}} \right)^2 \\ &\leq \left(\sum_{i_1,i_2,\ldots,i_d} f(i_1,i_2,\ldots,i_d)^2 \right)^{m_{\{j_1\},J_2}}, \end{aligned}$$

where we also used the fact that $m_{\{j_1\},J_2} \geq 2$.

Config. 8: $\left(\sum_{J_2} f(J_2)\right)^2$ with $|J_2| = k$, i.e. $l = 0$ in (34). Write $J_2 = \{j_1\} \cup J_2'$ and so $f(J_2) = f(\{j_1\}, J_2')f(J_2')$. Then split up J_2' into $J_2' = J_{2,1} \cup J_{2,2} \cup J_{2,3}$ with

$$\begin{aligned} J_{2,1} &= \{j_i \in J_2' \text{ that appears in } f(\{j_1\}, J_2') \text{ but not in } f(J_2')\}, \\ J_{2,2} &= \{j_i \in J_2' \text{ that appears both in } f(\{j_1\}, J_2') \text{ and in } f(J_2')\}, \\ J_{2,3} &= \{j_i \in J_2' \text{ that appears in } f(J_2') \text{ but not in } f(\{j_1\}, J_2')\}. \end{aligned}$$

Therefore

$$\sum_{J_2} f(J_2) =$$
$$\sum_{J_{2,2}} \Bigg[\sum_{j_1, J_{2,1}} f(\{j_1\}, J_{2,1}) f(\{j_1\}, J_{2,2}) f(\{j_1\}, J_{2,1}, J_{2,2}) \Bigg] \Bigg[f(J_{2,2}) \sum_{J_{2,3}} f(J_{2,2}, J_{2,3}) f(J_{2,3}) \Bigg].$$

If $J_{2,2} = \emptyset$, the situation reduces to

$$\left(\sum_{J_2} f(J_2) \right)^2 = \left(\sum_{j_1, J_{2,1}} f(j_1, J_{2,1}) \right)^2 \cdot \left(\sum_{J_{2,3}} f(J_{2,3}) \right)^2,$$

so that we can easily conclude with Config. 7 for the first term and by induction on k for the second one ($|J_{2,3}| \leq |J_2'| = k - 1$). If $J_{2,2} \neq \emptyset$, then by Cauchy-Schwarz inequality

$$\begin{aligned} &\left(\sum_{J_{2,2}} \Bigg[\sum_{j_1, J_{2,1}} f(j_1, J_{2,1}) f(j_1, J_{2,2}) \Bigg] \Bigg[f(J_{2,2}) \sum_{J_{2,3}} f(J_{2,2}, J_{2,3}) f(J_{2,3}) \Bigg] \right)^2 \\ &\qquad \leq \Bigg[\sum_{J_{2,2}} \Bigg(\sum_{j_1, J_{2,1}} f(j_1, J_{2,1}) f(j_1, J_{2,1}, J_{2,2}) f(j_1, J_{2,2}) \Bigg)^2 \Bigg] \\ &\qquad\qquad \times \Bigg[\sum_{J_{2,2}} f(J_{2,2})^2 \Bigg(\sum_{J_{2,3}} f(J_{2,2}, J_{2,3}) f(J_{2,3}) \Bigg)^2 \Bigg]. \end{aligned}$$

The two brackets in the latter bound can now be tackled with Config. 6, which achieves the induction procedure on k (and on d as well).

□

Proof of Proposition 4.1 in the general case. In the previous proof, it is easy to see that the symmetry assumption on f has only been used during the following steps: (1) in the initialization of the induction procedure on k, i.e. when $k = d$; (2) when studying Config. 1 and 7, i.e., each time the induction hypothesis on d is required. Therefore, these are the only situations we must go back to in order to extend the result to a non-fully-symmetric f.

Initialization step: $k = d$**.** Two kinds of configurations can occur here:

$$\sum_{j_1, \dots, j_l} \left(\sum_{j_{l+1}, \dots, j_d} f(\{j_1, \dots, j_d\}) \right)^2 \tag{45}$$

with $0 \leq l \leq d - 1$, and

$$\sum_{j_1, \dots, j_d} f(\{j_1, \dots, j_d\})^2. \tag{46}$$

As far as (45) is concerned, let us first observe that

$$\sum_{j_1,\dots,j_l}\Big(\sum_{j_{l+1},\dots,j_d} f(\{j_1,\dots,j_d\})\Big)^2 \;\le\; \Big(\sum_{j_1,\dots,j_d} |f(\{j_1,\dots,j_d\})|\Big)^2$$

$$\le\; \Big(\sum_{j_1,\dots,j_d}\prod_{\sigma\in\mathfrak{S}_d} |f(j_{\sigma(1)},\dots,j_{\sigma(d)})|^{m_{\sigma(1),\dots,\sigma(d)}}\Big)^2,$$

and we know that $\sum_{\sigma\in\mathfrak{S}_d} m_{\sigma(1),\dots,\sigma(d)} \ge 2$. Now, assume that there exists $\sigma_1\in\mathfrak{S}_d$ such that $m_{\sigma_1(1),\dots,\sigma_1(d)}\ge 2$ and $m_{\sigma(1),\dots,\sigma(d)}=0$ for $\sigma\neq\sigma_1$. Then obviously

$$\Big(\sum_{j_1,\dots,j_d} |f(j_{\sigma_1(1),\dots,\sigma_1(d)})|^{m_{\sigma_1(1),\dots,\sigma_1(d)}}\Big)^2 \le \Big(\sum_{j_1,\dots,j_d} f(j_1,\dots,j_d)^2\Big)^{m_{\sigma_1(1),\dots,\sigma_1(d)}},$$

which corresponds to the expected bound. Otherwise, there exists $\sigma_1\neq\sigma_2\in\mathfrak{S}_d$ such that $m_{\sigma_1(1),\dots,\sigma_1(d)}\ge 1$ and $m_{\sigma_2(1),\dots,\sigma_2(d)}\ge 1$. Then, by using Cauchy-Schwarz, we get

$$\begin{aligned}
&\Big(\sum_{j_1,\dots,j_d}\Big[|f(j_{\sigma_1(1)},\dots,j_{\sigma_1(d)})|^{m_{\sigma_1(1),\dots,\sigma_1(d)}}\Big]\\
&\qquad\times\Big[|f(j_{\sigma_2(1)},\dots,j_{\sigma_2(d)})|^{m_{\sigma_2(1),\dots,\sigma_2(d)}}\prod_{\sigma\in\mathfrak{S}_d\setminus\{\sigma_1,\sigma_2\}}|f(j_{\sigma(1)},\dots,j_{\sigma(d)})|^{m_{\sigma(1),\dots,\sigma(d)}}\Big]\Big)^2\\
&\qquad\qquad\le\Big[\sum_{j_1,\dots,j_d}|f(j_{\sigma_1(1)},\dots,j_{\sigma_1(d)})|^{2m_{\sigma_1(1),\dots,\sigma_1(d)}}\Big]\\
&\times\Big[\sum_{j_1,\dots,j_d}|f(j_{\sigma_2(1)},\dots,j_{\sigma_2(d)})|^{2m_{\sigma_2(1),\dots,\sigma_2(d)}}\prod_{\sigma\in\mathfrak{S}_d\setminus\{\sigma_1,\sigma_2\}}|f(j_{\sigma(1)},\dots,j_{\sigma(d)})|^{2m_{\sigma(1),\dots,\sigma(d)}}\Big].
\end{aligned}\tag{47}$$

Of course,

$$\sum_{j_1,\dots,j_d}|f(j_{\sigma_1(1)},\dots,j_{\sigma_1(d)})|^{2m_{\sigma_1(1),\dots,\sigma_1(d)}}\le\Big(\sum_{j_1,\dots,j_d} f(j_1,\dots,j_d)^2\Big)^{m_{\sigma_1(1),\dots,\sigma_1(d)}}.$$

At this point, if $m_{\sigma(1),\dots,\sigma(d)}=0$ for every $\sigma\notin\{\sigma_1,\sigma_2\}$, the proof is concluded thanks to (47). Otherwise, there exists $\sigma_3\notin\{\sigma_1,\sigma_2\}$ such that $m_{\sigma_3(1),\dots,\sigma_3(d)}\ge 1$ and we have (use the trivial estimate $\sum_i a_ib_i\le(\sum_i a_i)(\sum_i b_i)$ for $a_i,b_i\ge 0$)

$$\begin{aligned}
&\sum_{j_1,\dots,j_d}|f(j_{\sigma_2(1)},\dots,j_{\sigma_2(d)})|^{2m_{\sigma_2(1),\dots,\sigma_2(d)}}\prod_{\sigma\in\mathfrak{S}_d\setminus\{\sigma_1,\sigma_2\}}|f(j_{\sigma(1)},\dots,j_{\sigma(d)})|^{2m_{\sigma(1),\dots,\sigma(d)}}\\
&\le\Big[\sum_{j_1,\dots,j_d}|f(j_{\sigma_2(1)},\dots,j_{\sigma_2(d)})|^{2m_{\sigma_2(1),\dots,\sigma_2(d)}}\Big]\times\\
&\quad\Big[\sum_{j_1,\dots,j_d}|f(j_{\sigma_3(1)},\dots,j_{\sigma_3(d)})|^{2m_{\sigma_3(1),\dots,\sigma_3(d)}}\prod_{\sigma\in\mathfrak{S}_d\setminus\{\sigma_1,\sigma_2,\sigma_3\}}|f(j_{\sigma(1)},\dots,j_{\sigma(d)})|^{2m_{\sigma(1),\dots,\sigma(d)}}\Big]\\
&\le\Big(\sum_{j_1,\dots,j_d} f(j_1,\dots,j_d)^2\Big)^{m_{\sigma_2(1),\dots,\sigma_2(d)}}\times\\
&\quad\Big[\sum_{j_1,\dots,j_d}|f(j_{\sigma_3(1)},\dots,j_{\sigma_3(d)})|^{2m_{\sigma_3(1),\dots,\sigma_3(d)}}\prod_{\sigma\in\mathfrak{S}_d\setminus\{\sigma_1,\sigma_2,\sigma_3\}}|f(j_{\sigma(1)},\dots,j_{\sigma(d)})|^{2m_{\sigma(1),\dots,\sigma(d)}}\Big].
\end{aligned}$$

We can now repeat the procedure to handle the latter bracket. The treatment of (46) can clearly be derived from the same argument.

Back to Config. 1: $\sum_{j_1,J_1} f(\{j_1\},J_1)^2\Big(\sum_{J_2} f(\{j_1\},J_1,J_2)f(\{j_1\},J_2)\Big)^2$ with $|J_1|+|J_2| = k-1$. As each of the terms here contains the index j_1, it seems natural to use the same idea as in the fully-symmetric case, that is, to "freeze" j_1 so to conclude by induction on d. The problem in the non-fully-symmetric case lies of course in the fact that j_1 may appear in any position of $f(j_{i_1},\dots,j_{i_d})$. We are therefore led to introduce the functions $f_j^{(p)} : \{1,\dots,N\}^{d-1} \to \mathbb{R}$ ($p \in \{1,\dots,d\}, j \in \{1,\dots,N\}$) defined as

$$f_j^{(p)}(j_1,\dots,j_{d-1}) := f(j_1,\dots,j_{p-1},j,j_p,\dots,j_{d-1}).$$

With the notation used in this section, we have

$$f(\{j_1\},J_1) = f_{j_1}^{(1)}(J_1)\dots f_{j_1}^{(d)}(J_1)$$

where each $f_{j_1}^{(p)}(J_1)$ is associated with the weights

$$m^{(p)}_{i_1,\dots,i_{d-1}} := m_{i_1,\dots,i_{p-1},1,i_p,\dots,i_{d-1}}.$$

Thus,

$$\begin{aligned}&\sum_{j_1,J_1} f(\{j_1\},J_1)^2\Big(\sum_{J_2} f(\{j_1\},J_1,J_2)f(\{j_1\},J_2)\Big)^2\\ &= \sum_{j_1,J_1}\Big[f_{j_1}^{(1)}(J_1)^2\dots f_{j_1}^{(d)}(J_1)^2\Big]\Big(\sum_{J_2}\Big[\{f_{j_1}^{(1)}(J_1,J_2)f_{j_1}^{(1)}(J_2)\}\dots\{f_{j_1}^{(d)}(J_1,J_2)f_{j_1}^{(d)}(J_2)\}\Big]\Big)^2. \quad (48)\end{aligned}$$

We denote by $1 \le p_1 < \dots < p_r \le d$ (resp. $1 \le q_1 < \dots < q_s \le d$) the positions of j_1 that indeed come out in the first (resp. second) bracket of (48), i.e., such that

$$m_{J_1}^{(p_i)} \ge 1 \qquad (\text{resp. } \; m_{J_1,J_2}^{(q_i)} + m_{J_2}^{(q_i)} \ge 1).$$

The sum (48) can accordingly be written as

$$\sum_{j_1,J_1}\Big[f_{j_1}^{(p_1)}(J_1)^2\dots f_{j_1}^{(p_r)}(J_1)^2\Big]\Big(\sum_{J_2}\Big[\{f_{j_1}^{(q_1)}(J_1,J_2)f_{j_1}^{(q_1)}(J_1)\}\dots\{f_{j_1}^{(q_s)}(J_1,J_2)f_{j_1}^{(q_s)}(J_1)\}\Big]\Big)^2. \quad (49)$$

Let us now consider the two situations where $s = 0$ and $s \ge 1$ separately.

Sub-Config. 1.1: $s = 0$. The sum reduces here to

$$\sum_{j_1,J_1}\big[f_{j_1}^{(p_1)}(J_1)^2\big]\big[f_{j_1}^{(p_2)}(J_1)^2\dots f_{j_1}^{(p_r)}(J_1)^2\big]. \quad (50)$$

If $r = 1$, then the second bracket (50) is equal to 1 and we are in a position to apply the induction hypothesis (on d) to $f_{j_1}^{(p_1)}$, which entails

$$\begin{aligned}\sum_{j_1,J_1} f_{j_1}^{(p_1)}(J_1)^2 &\le \sum_{j_1}\Big(\sum_{i_2,\dots,i_d} f_{j_1}^{(p_1)}(i_2,\dots,i_d)^2\Big)^{m_{J_1}^{(p_1)}}\\ &\le \Big(\sum_{j_1,i_2,\dots,i_d} f_{j_1}^{(p_1)}(i_2,\dots,i_d)^2\Big)^{m_{J_1}^{(p_1)}} \le \Big(\sum_{i_1,\dots,i_d} f(i_1,\dots,i_d)^2\Big)^{m_{\{j_1\},J_1}}\end{aligned}$$

since in this case $m^{(p_1)}_{J_1} = m_{\{j_1\},J_1} \geq 1$. If $r \geq 2$ in (50), divide J_1 into $J_1 = J_{1,1} \cup J_{1,2} \cup J_{1,3}$ with

$$\begin{aligned} J_{1,1} &:= \{j_i \in J_1 \text{ that appears in the first bracket of (50) but not in the second one}\}, \\ J_{1,2} &:= \{j_i \in J_1 \text{ that appears in both brackets of (50)}\}, \\ J_{1,3} &:= \{j_i \in J_1 \text{ that appears in the second bracket of (50) but not in the first one}\}. \end{aligned}$$

So, (50) can be written as

$$\begin{aligned} &\sum_{j_1,J_{1,2}} \left[\sum_{J_{1,1}} f^{(p_1)}_{j_1}(J_{1,1} \cup J_{1,2})^2 \right] \left[\sum_{J_{1,3}} f^{(p_2)}_{j_1}(J_{1,2} \cup J_{1,3})^2 \dots f^{(p_r)}_{j_1}(J_{1,2} \cup J_{1,3})^2 \right] \\ &\qquad \leq \left[\sum_{j_1,J_{1,1},J_{1,2}} f^{(p_1)}_{j_1}(J_{1,1} \cup J_{1,2})^2 \right] \left[\sum_{j_1,J_{1,2},J_{1,3}} f^{(p_2)}_{j_1}(J_{1,2} \cup J_{1,3})^2 \dots f^{(p_r)}_{j_1}(J_{1,2} \cup J_{1,3})^2 \right]. \end{aligned} \tag{51}$$

By setting $J_1' := J_{1,1} \cup J_{1,2}$, we get as above

$$\sum_{j_1,J_1'} f^{(p_1)}_{j_1}(J_1')^2 = \sum_{j_1} \left(\sum_{J_1'} f^{(p_1)}_{j_1}(J_1')^2 \right) \leq \left(\sum_{i_1,\dots,i_d} f(i_1,\dots,i_d)^2 \right)^{m^{(p_1)}_{J_1'}}.$$

Note that by the very definition of $J_{1,1}, J_{1,2}$, we have $m^{(p_1)}_{J_1'} = \sum_{1 \leq i_1,\dots,i_{d-1} \leq k} m^{(p_1)}_{i_1,\dots,i_{d-1}}$. We now easily repeat the procedure (i.e., we go back to the step (50)) to deal with the second bracket of (51).

Sub-Config. 1.2: $s \geq 1$. In (49), divide J_1 into $J_1 = J_{1,1} \cup J_{1,2} \cup J_{1,3}$ with

$$\begin{aligned} J_{1,1} &:= \{j_i \in J_1 \text{ that appears in the first bracket but not in the second one}\}, \\ J_{1,2} &:= \{j_i \in J_1 \text{ that appears in both brackets}\}, \\ J_{1,3} &:= \{j_i \in J_1 \text{ that appears in the second bracket but not in the first one}\}. \end{aligned}$$

Quantity (49) becomes

$$\begin{aligned} &\sum_{j_1,J_{1,2}} \left[\sum_{J_{1,1}} \prod_{1 \leq i \leq r} f^{(p_i)}_{j_1}(J_{1,1} \cup J_{1,2})^2 \right] \left[\sum_{J_{1,3}} \left(\sum_{J_2} \prod_{1 \leq l \leq s} f^{(q_l)}_{j_1}(J_{1,2} \cup J_{1,3}, J_2) f^{(q_l)}_{j_1}(J_{1,2} \cup J_{1,3}) \right)^2 \right] \\ &\qquad \leq \left[\sum_{j_1,J_{1,1},J_{1,2}} \dots \right] \left[\sum_{j_1,J_{1,2},J_{1,3}} \dots \right], \end{aligned} \tag{52}$$

where we have only used the crude bound $\sum_i a_i b_i \leq (\sum_i a_i)(\sum_i b_i)$ for $a_i, b_i \geq 0$. The first bracket in the bound (52) corresponds to Sub-Config 1.1. As for the second bracket, it can be written, if $J_1' := J_{1,2} \cup J_{1,3}$, as

$$\sum_{j_1,J_1'} \left(\sum_{J_2} \left[f^{(q_1)}_{j_1}(J_1', J_2) f^{(q_1)}_{j_1}(J_2) \right] \left[f^{(q_2)}_{j_1}(J_1', J_2) f^{(q_2)}_{j_1}(J_2) \dots f^{(q_s)}_{j_1}(J_1', J_2) f^{(q_s)}_{j_1}(J_2) \right] \right)^2. \tag{53}$$

At this point, if $s = 1$, the result is easily achieved by applying the induction hypothesis (on d) to $f^{(q_1)}_{j_1}$. If $s \geq 2$, split up J_2 into $J_2 = J_{2,1} \cup J_{2,2} \cup J_{2,3}$ with

$$\begin{aligned} J_{2,1} &:= \{j_i \in J_2 \text{ that appears in the first bracket of (53) but not in the second one}\}, \\ J_{2,2} &:= \{j_i \in J_2 \text{ that appears in both brackets of (53)}\}, \\ J_{2,3} &:= \{j_i \in J_2 \text{ that appears in the second bracket of (53) but not in the first one}\}. \end{aligned}$$

With this splitting in hand, we appeal to the (now) usual arguments: apply Cauchy-Schwarz inequality over $J_{2,2}$, then use an appropriate division for J_1' (just as in (50)) and finally resort to

the induction hypothesis on d to estimate the expression involving $f_{j_1}^{(q_1)}$ only. In this way, it is not difficult to realize that the procedure reduces to the estimation of

$$\sum_{j_1,J_1''}[f_{j_1}^{(q_2)}(J_1'')^2\dots f_{j_1}^{(q_s)}(J_1'')^2]\Big(\sum_{J_2''}[f_{j_1}^{(q_2)}(J_1'',J_2'')f_{j_1}^{(q_2)}(J_1'')]\dots[f_{j_1}^{(q_s)}(J_1'',J_2'')f_{j_1}^{(q_s)}(J_1'')]\Big)^2$$

for two sets of indices J_1'', J_2'' that verify Condition (33). In other words, we are brought back to the configuration (49), with this time $s-1$ positions of j_1 involved in the second sum (over J_2''). As a consequence, it suffices to repeat the reasoning $s-1$ times.

Back to Config. 7: $\Big(\sum_{j_1,J_2} f(\{j_1\},J_2)\Big)^2$ with $|J_2|=k-1$. Bear the previous notation in mind. One has

$$\Big(\sum_{j_1,J_2} f(\{j_1\},J_2)\Big)^2=\Big(\sum_{j_1,J_2}\Big[f_{j_1}^{(p_1)}(J_2)\Big]\Big[f_{j_1}^{(p_2)}(J_2)\dots f_{j_1}^{(p_s)}(J_2)\Big]\Big)^2, \tag{54}$$

where the positions $1\le p_1<\dots<p_s\le d$ are such that $m_{J_2}^{(p_i)}\ge 1$. If $s=1$, the conclusion is easily derived by induction on d (just as in the fully-symmetric case). If $s\ge 2$, then divide J_2 into $J_2=J_{2,1}\cup J_{2,2}\cup J_{2,3}$ with

$$\begin{aligned}
J_{2,1} &:= \{j_i\in J_2 \text{ that appears in the first bracket of (54) but not in the second one}\},\\
J_{2,2} &:= \{j_i\in J_2 \text{ that appears in both brackets of (54)}\},\\
J_{2,3} &:= \{j_i\in J_2 \text{ that appears in the second bracket of (54) but not in the first one}\}.
\end{aligned}$$

Write now (54) as

$$\begin{aligned}
&\Big(\sum_{j_1,J_{2,2}}\Big[f_{j_1}^{(p_1)}(J_{2,2})\sum_{J_{2,1}}f_{j_1}^{(p_1)}(J_{2,1})f_{j_1}^{(p_1)}(J_{2,1},J_{2,2})\Big]\times\\
&\qquad\Big[\prod_{2\le i\le r}f_{j_1}^{(p_i)}(J_{2,2})\sum_{J_{2,3}}\prod_{2\le i\le r}f_{j_1}^{(p_i)}(J_{2,3})f_{j_1}^{(p_i)}(J_{2,2},J_{2,3})\Big]\Big)^2\\
&\le\Big[\sum_{j_1,J_{2,2}}f_{j_1}^{(p_1)}(J_{2,2})^2\Big(\sum_{J_{2,1}}f_{j_1}^{(p_1)}(J_{2,1})f_{j_1}^{(p_1)}(J_{2,1},J_{2,2})\Big)^2\Big]\\
&\qquad\Big[\sum_{j_1,J_{2,2}}\prod_{2\le i\le r}f_{j_1}^{(p_i)}(J_{2,2})^2\Big(\sum_{J_{2,3}}\prod_{2\le i\le r}f_{j_1}^{(p_i)}(J_{2,3})f_{j_1}^{(p_i)}(J_{2,2},J_{2,3})\Big)^2\Big],
\end{aligned}$$

which brings us back to Config. 1 and concludes the proof. □

Acknowledgement. We are grateful to Todd Kemp and Roland Speicher for helpful comments and references about hypercontractivity.

References

[1] P. Biane and R. Speicher (1998). Stochastic analysis with respect to free Brownian motion and analysis on Wigner space. *Probab. Theory Rel. Fields* **112**, 373-409.

[2] A. Deya and I. Nourdin (2011). Convergence of Wigner integrals to the tetilla law. *ALEA*, in revision.

[3] V. Kargin (2007). A proof of a non-commutative central limit theorem by the Lindeberg method. *Elect. Comm. in Probab.* **12**, 36-50.

[4] T. Kemp, I. Nourdin, G. Peccati and R. Speicher (2011). Wigner chaos and the fourth moment. *Ann. Probab.*, in press.

[5] J.W. Lindeberg (1922). Eine neue Herleitung des exponential-Gesetzes in der Warscheinlichkeitsrechnung. *Math. Zeitschrift.* **15**, 211-235.

[6] E. Mossel, R. O'Donnell and K. Oleszkiewicz (2010). Noise stability of functions with low influences: invariance and optimality. *Ann. Math.* **171**, no. 1, 295-341.

[7] A. Nica and R. Speicher (1998). Commutators of free random variables. *Duke Math. J.* **92**, no. 3, 553-592.

[8] A. Nica and R. Speicher (2006). *Lectures on the Combinatorics of Free Probability*. Cambridge University Press.

[9] I. Nourdin (2011). Yet another proof of the Nualart-Peccati criterion. *Electron. Comm. Probab.* **16**, 467-481.

[10] I. Nourdin, G. Peccati and G. Reinert (2010). Invariance principles for homogeneous sums: universality of Gaussian Wiener chaos. *Ann. Probab.* **38**, no. 5, 1947-1985.

[11] D. Nualart and G. Peccati (2005). Central limit theorems for sequences of multiple stochastic integrals. *Ann. Probab.* **33** (1), 177-193.

[12] G. Peccati and C. Zheng (2011). Universal Gaussian fluctuations on the discrete Poisson chaos. Preprint.

[13] D.V. Voiculescu (1985). Symmetries of some reduced free product C^*-algebras. *Operator algebras and their connection with topology and ergodic theory*, Springer Lecture Notes in Mathematics **1132**, 556–588.